N.º duplo 3/4 - 88 - ABRIL 1987 - Esc. 300$00

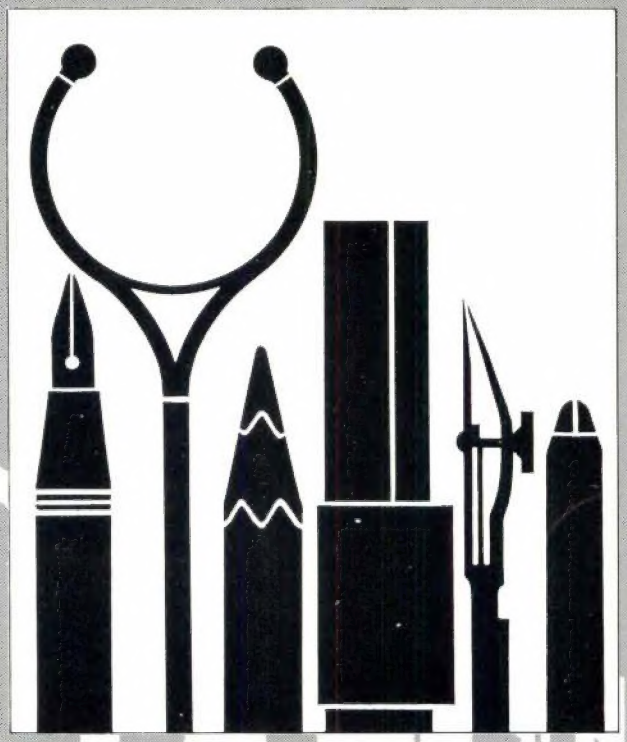
técnica

Revista de Engenharia
Publicação da Associação dos Estudantes do Instituto Superior Técnico

A.E.I.S.T. - 75 ANOS

Financiamento
Investimento das
PROFISSÕES LIBERAIS

O SEU PONTO DE PARTIDA



A sua profissão?
Médico, engenheiro, arquitecto, economista, advogado ...
Uma actividade que exige iniciativa, dinamismo, investimento.
Uma carreira em que cada dia é um novo ponto de partida.
O Financiamento ao Investimento das Profissões Liberais é uma Linha de Crédito do Montepio Geral que dá um apoio seguro à sua actividade, na compra de bens imóveis, fracções de prédios ou lojas, equipamento médico, informático ou outros equipamentos específicos.
O financiamento ao Investimento das Profissões Liberais é um novo ponto de partida, para si, que quer ir mais longe.
Consulte os serviços do Montepio Geral para informações mais detalhadas.

MONTEPIO GERAL
CAIXA ECONOMICA DE LISBOA

APOIO DINÂMICO A GENTE DINÂMICA

CIESA MCK

PROPRIEDADE DA ASSOCIAÇÃO DOS ESTUDANTES DO INSTITUTO SUPERIOR TÉCNICO • **DIRECTOR:** M. DE ABREU FARO • **COORDENADOR EDITORIAL:** J. ANTUNES BARATA • **DIRECÇÃO E ADMINISTRAÇÃO:** AEIST - Av. Rovisco Pais, 1000 LISBOA - Tels.: 88 93 23/88 10 18 • **PUBLICIDADE** TÉCNICA - AEIST - Av. Rovisco Pais, 1000 LISBOA - Tels.: 88 10 18/89 75 15 • **COMPOSIÇÃO E IMPRESSÃO:** CABOGRÁFICA, LDA., Rua das Biscoiteiras, Lote 18 (traseiras) - 2795 LINDA-A-VELHA • PUBLICAÇÃO TRIMESTRAL - 5000 EXEMPLARES • ASSINATURAS (4 números): PORTUGAL 500$00 - ESTRANGEIRO 12US$

técnica

3/4 • 86

ABRIL 1987

3 EDITORIAL: A Associação dos Estudantes do Instituto Superior Técnico, a AEIST, foi fundada em 11 de Dezembro de 1911

4 Palavras proferidas na Sessão de Encerramento das Comemorações dos 75 anos da Associação de Estudantes do I.S.T. por:
A. Simões Lopes, Vice-Reitor no exercício das funções de Reitor da UTL
Abreu Faro, Professor do I.S.T.
Francisco Sanches, da Direcção da AEIST
Venâncio Rosa, da Comissão Organizadora das Comemorações

ARTIGOS CONVIDADOS

15 FÍSICA DE ALTAS ENERGIAS: Actividade recente do Grupo de Física Experimental de Altas Energias de Lisboa
José Mariano Gago, Amélia Maio, Mário Pimenta, João Varela

24 TEORIA E PROPRIEDADES DE PLASMAS PRODUZIDOS POR ONDAS DE SUPERFÍCIE
Carlos Matos Ferreira

40 USO DE MODELOS DE MARKOV NÃO OBSERVÁVEIS NO RECONHECIMENTO DE PALAVRAS ISOLADAS
António Joaquim Serralheiro, José Manuel S. Tribolet

ARTIGOS

52 COLECÇÃO DE PROGRAMAS E FILTROS DESTINADOS A OPERAR SOBRE "CONJUNTOS VAGOS"
A. J. N. Martins Aguiar

75 OS ZEÓLITOS E SUAS APLICAÇÕES INDUSTRIAIS
F. Ramôa Ribeiro, F. Lemos

NOTAS CIENTIFICAS

81 A TRANSMISSÃO DO CALOR EM RELATIVIDADE
António Brotas

86 VIABILIDADE ECONÓMICA DE BATERIAS DE LÍTIO PARA VEÍCULOS ELÉCTRICOS
C. A. C. de Sequeira

89 ON THE CALCULATION OF THE FOUR-VECTOR (k, ω/c) IN A LINEAR TRANSFORMATION OF COORDINATES, AN EXAMPLE
Maria Hermínia Marçal

93 FORMATION OF DIFFUSION COATINGS ON IRON AND STEEL. PARTE I: FORMATION PROCESSES
C. A. C. Sequeira and **C. M. G. S. Nunes**

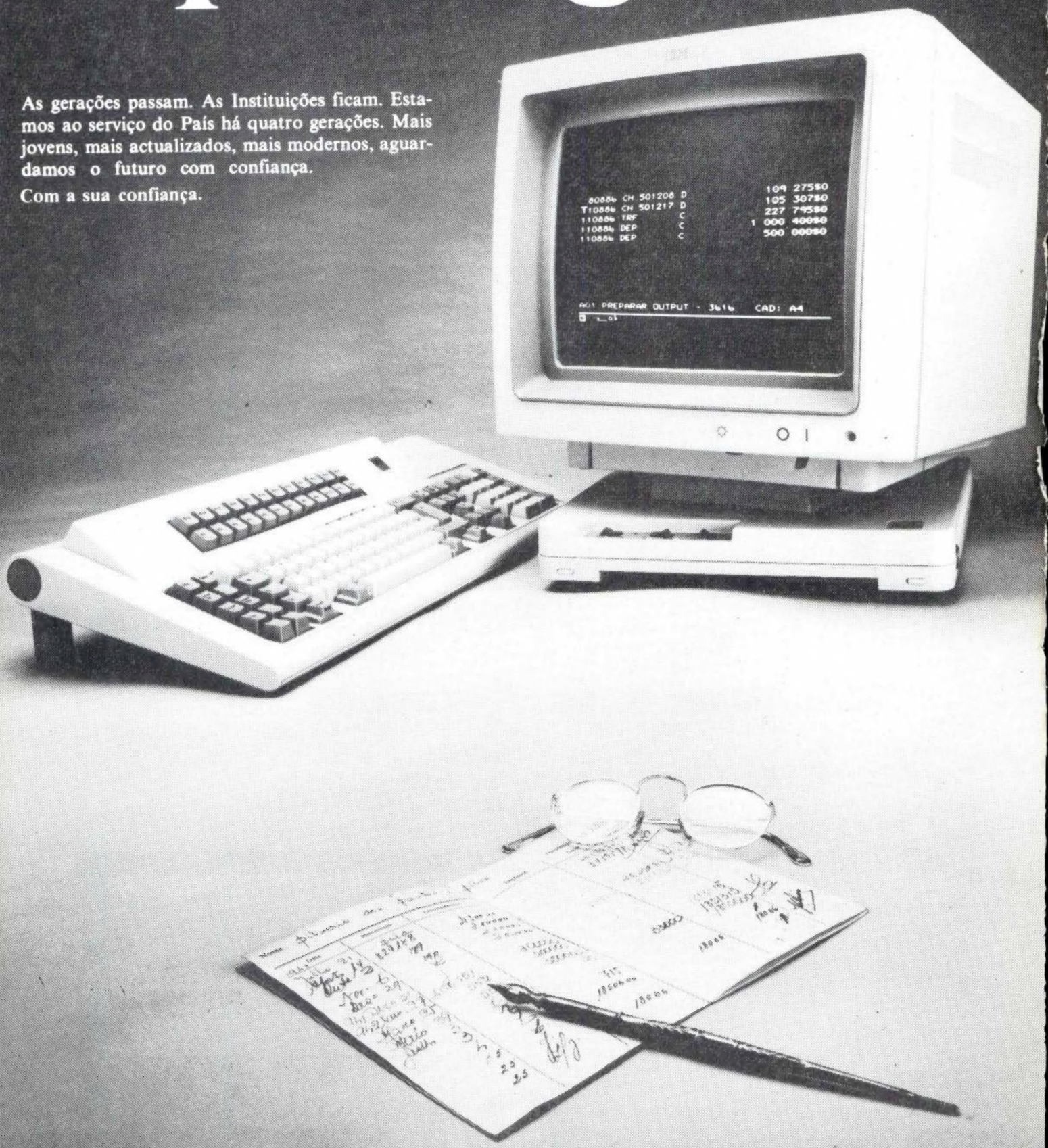
100 THE DISPERSION EQUATION FOR ELECTROSTATIC WAVES IN A BEAM-PLASMA SYSTEM IMMERSED IN AN INHOMOGENEOUS MAGNETIC FIELD
C. A. F. Varandas and **J. A. C. Cabral**

ISSN 0040-1714

A quarta geração

As gerações passam. As Instituições ficam. Estamos ao serviço do País há quatro gerações. Mais jovens, mais actualizados, mais modernos, aguardamos o futuro com confiança.

Com a sua confiança.



CAIXA GERAL DE DEPÓSITOS

A Associação dos Estudantes do Instituto Superior Técnico, a AEIST, foi fundada em 11 de Dezembro de 1911

No tempo, 75 anos que são passados,
a fundação do Instituto e a fundação da Associação coincidem.
Praticamente são factos que emergem na mesma data e na mesma atitute.
Prenúncio de que o Instituto e a sua Associação de Estudantes
ficariam ligados.

De facto assim tem sido.
A Associação nasceu do Técnico,
mas o Técnico, o Instituto, deve parte importante da sua configuração e
prestígio ao que a sua Associação de Estudantes tem sabido e
desejado ser.

Na sessão de 11 de Dezembro comemorou-se a fundação da Associação dos Estudantes do I.S.T.
Este número da TÉCNICA é especialmente dedicado a essa efeméride.
Quanto se disse nessa cerimónia aqui fica para sinal e reflexão.

O belo retrato de Alfredo Bensaúde, do número anterior, parece que nos chega de um passado distante.
Medem-se os 75 anos do I.S.T. e da A.E.I.S.T. no ar antigo desse retrato.
No entanto, as notas Histórico-Pedagógicas são de modernidade que ainda não se cumpriu.

Neste número, o diamante
símbolo das Bodas de Diamante
invade com a sua luz o futuro.
Nesse futuro, como no passado, tudo continuará a ser trivial e simples:

É estudante quem estudar.
É professor quem ensinar.
Que possa estudar quem estuda e que possa ensinar
quem saiba e se ofereça.

Que o Instituto Superior Técnico e a Associação dos Estudantes do I.S.T., unidos e em preocupação comum, cuidem destes factos simples e que exactamente por o serem definirão o futuro.

Manuel José de Abreu Faro

Palavras proferidas na Sessão de encerramento das comemorações dos 75 anos da Associação de Estudantes do I.S.T. por

A. SIMÕES LOPES
Vice-Reitor no exercício das funções de Reitor da UTL

Direi que é para mim extremamente grato aproveitar a oportunidade para saudar os alunos do Instituto Superior Técnico através da sua Associação; para saudar essa mesma Associação que acaba de completar 75 anos de existência e para, em reflexão breve, formular votos de muito empenhamento e de grande sucesso a todos quantos vêm dedicando parte significativa das suas energias ao movimento associativo, com intenções de organizar os estudantes para os grandes projectos do futuro.

Sendo professor de outra Instituição da Universidade, sempre tenho apreciado, afirmado e defendido do Instituto Superior Técnico a imagem de uma grande Escola da Universidade Técnica e do país, qualquer que seja o índice que se utilize na aferição da grandeza. Essa imagem é fruto de um longo e persistente trabalho de mais de 75 anos, em que se vem tornando hábito referenciar figuras-chaves do passado, desde Alfredo Bensaúde e Duarte Pacheco, professores prestigiosos que muito prestigiaram a Escola, havendo certa inibição, em referir os presentes, com o receio natural de a algum esquecer e com a certeza de que a muitos feriríamos na modéstia que anda normalmente de par com as qualidades positivas do cientista, do universitário, do Homem. Mas são os professores os mais frequentemente referidos, por razões óbvias, sempre se secundarizando aparentemente os alunos, nesse papel de acumulação do prestígio da Instituição.

Não me consta que tenha havido, até hoje, contestação a esta atitude; e louvável (e correcto) é que não exista, não só porque nos velhos Mestres (e nos bons Mestres do presente) os alunos sempre sabem distinguir a imagem séria e digna dos que trabalham para a causa institucional, mas também porque neles se perfilam modelos de estudante, que o têm de ser sempre, porque a qualidade, a seriedade, a dignidade da docência exigirão em pleno, em permanência, a preparação, o trabalho, o *studio*. Não pode conceber-se, muito menos aceitar-se, o professor que não estuda; o professor é antes do mais, e sempre, um estudante, até para merecer a consideração e o apreço dos *alumni* que lhe são confiados e se lhe confiam.

Mas as reflexões que quero aproveitar para fazer vão noutra direcção, e penso que se justificam sem reservas. Quero deixá-las à consideração de todos, neste dia que é de festa e de celebração, mas em que olhar para o futuro não é menos importante do que glorificar o passado.

Estará perfeitamente criado um verdadeiro espírito de comunidade, nesta grande Escola de cerca de 6000 alunos? Estaremos todos a dar o nosso melhor esforço no sentido de instalar, em bases perfeitamente sólidas, um amplo espírito de comunidade nesta nossa Universidade onde o número de alunos anda perto dos 15 000? E será importante criar-se esse espírito comunitário?

Tenho-me dado conta, mais vezes do que desejaria, de que uma certa compartimentação existe entre os alunos desta Escola, mais do que em qualquer outra da Universidade, o que de certo, tem a ver com o tamanho da Instituição. O exemplo mais frisante surge-me a nível dos finalistas, sobretudo, quando começam a organizar as suas viagens de fim de curso para que se vai tornando habitual dirigirem-se à Reitoria para apoio — tantas vezes, infelizmente, para receberem apoio "moral" apenas; e a observação que faço, regularmente, é a de que cada Departamento elabora o seu programa com grande alheamento do programa dos colegas de outros Departamentos, por vezes, mesmo, com largo desconhecimento deles.

Direi que não é esse alheamento causa do frouxo apoio que a Reitoria pode dar; isto é, o apoio

talvez não fosse significativamente diferente se alguma coordenação houvesse entre os programas elaborados. Mas, a mim, causa-me apreensão sincera o facto de a compartimentação existir, porque isso quer dizer que o convívio é menos intenso e diversificado do que eu desejaria (e se desejaria). Donde, o não aproveitamento de todas as potencialidades formativas que na Escola existem, porque ao engenheiro mecânico de amanhã (de hoje) há-de em muitas circunstâncias assistir a necessidade do trabalho conjunto com o engenheiro químico, ou de minas,... que sei eu... amanhã, hoje mesmo... e, no entanto, virá a ser depois da Escola que há-de vencer-se a barreira para a actividade interdisciplinar, designadamente para o conhecimento dos Colegas dos Grupos que hão-de constituir-se.

E a nível da Universidade? Salvo casos que me recuso mesmo assim a não considerar esporádicos, até para minha defesa do conceito de Universidade e da sua operacionalidade, que conhecimento alargado terão os alunos do "Técnico" dos seus colegas, e do trabalho dos seus colegas, das outras Escolas? Será, por exemplo, que todos estarão seguros de conhecer quantas e quais são as outras Escolas da sua Universidade?

E aqui também vem instalar-se a angústia, porque de novo se deixa de aproveitar as potencialidades formativas existentes na Universidade, pela compartição excessiva que se instalou. Quantos alunos finalistas de engenharia civil, interessados no planeamento urbano, têm convivido com colegas seus da Faculdade de Arquitectura, do Instituto Superior de Agronomia (Arquitectura paisagista), do Instituto Superior de Economia (Economia Regional e Economia Urbano), do Instituto Superior de Ciências Sociais e Políticas (Sociologia Urbana e Sociologia Rural), interessados na mesma temática?

E, no entanto, a actividade de cada um é cada vez menos destacável, cada vez menos isolável da actividade dos outros...

Entretanto, falamos da necessidade de alargamento do conhecimento além-fronteiras, e defendêmo-lo, designadamente no sentido de aumentar a mobilidade dos estudantes que lhes permita viver e estudar noutros países da Comunidade Económica Europeia, como base (a mais sólida) de formação da Comunidade dos Cidadãos. Mas estará feito o essencial para nos conhecermos e organizarmos, em Portugal, como Comunidade?

Deixo estas reflexões a todos vós, porque embora sabendo que não é de vós que se espera a contribuição total para a solução do problema, a verdade é que ela em parte de vós que se espera a contribuição total para a solução do problema, a verdade é que ela em parte de vós depende, se a sensibilização existir e se a motivação despertar para um trabalho mais plural, mais conjunto, mais interligado. Sei que a Associação de Estudantes se empenha nesta perspectiva, em que tem actividade válida; e a testá-lo está a eleição recente para a Direcção da Associação Académica de Lisboa, que muito gostosamente justificou de mim os cumprimentos e os desejos de bom labor; mas será esta minha preocupação, esta preocupação da Direcção da vossa Associação, preocupação de todos?

Retomo, para terminar, as minhas considerações iniciais.

Encerrou-se, há semanas, o Ciclo de Comemorações dos 75 anos do Instituto Superior Técnico: foi motivo de alegria e de manifestação convicta de esperança no futuro desta grande Instituição. Termina hoje o Ciclo de Comemorações dos 75 anos da Associação de Estudantes do "Técnico": é motivo de não menor alegria nem menor esperança no futuro da Instituição — porque na juventude elas residem, a alegria, a esperança, a força, a generosidade, a persistência, que fazem as Obras Grandes.

Que seja grande a vossa Obra!

Palavras proferidas na Sessão de encerramento das comemorações dos 75 anos da Associação de Estudantes do I.S.T. por

ABREU FARO
Professor do I.S.T.

Julgo que Einstein nunca proferiu o lugar comum "tudo é relativo". Pelo menos no que respeita à Teoria da Relatividade.

Pelo contrário. O Princípio da Relatividade liberta-nos de subjectivismo. Para qualquer observador as leis da Física são as mesmas onde quer que esteja e como esteja. Até é possível escrevê-las sob a mesma forma.

Assim, e sem que haja contradição, a Relatividade é coisa bem diferente do relativismo de um Protágoras e, mais nos nossos dias, de um mesmo relativismo de Kant.

As Leis da Física, são da Física, mas tenho para mim que a verdade sendo verdade vai além do aspecto particular que reveste. Havendo humanização colhe-se luz das mais diversas fontes.

Vem isto a propósito de quem pode apurar a verdade.

Relaciona-se isto com o facto de quem e de que palavras nos podemos socorrer para testemunho.

Fiéis ao Princípio da Relatividade qualquer um o pode fazer.

No entanto, para que seja credível é necessário que ligue as palavras a juízos de realidade e de dignidade. É necessário que haja empenhamento, compromisso na afirmação de acções e resultados.

Pela naturalidade e justeza que lhes são inerentes o enunciado das verdades primeiras corre sempre o risco de não se afirmar como conselho, de tal modo se confunde a aceitação dessas verdades com a prática que desse facto, e em coerência, deveria decorrer. É um dizer, um saber incontroverso de quase nenhum efeito.

Isso, portanto, e infelizmente, não basta.

Com esses enunciados universais se diz sim e se diz não e quase sempre o indesejável "tudo é relativo".

É necessário que as palavras do testemunho se enriqueçam com a legitimidade da experiência ou do recurso impecável aos primeiros princípios.

Gozam destas propriedades as Notas Histórico-Pedagógicas sobre o Instituto Superior Técnico de Alfredo Bensaúde.

"Nas notas que se seguem, exponho as ideias directrizes que nos guiaram, descrevo o que se pretendeu realizar e até que ponto foi possível fazê-lo, comento os processos empregados no Instituto, em contraposição com os tradicionais nas nossas escolas, que pusemos de parte, e historio também algumas das dificuldades que foi preciso vencer para que a escola vingasse...

"Considero a redacção destas notas como o último acto da missão de que me encarregou o Governo Provisório..."

Que o observador pode ser qualquer prova-o o facto de muitas das passagens das Notas Histórico-Pedagógicas poderem ter sido escritas por alunos.

Omita-se o autor, saliente-se aqui e ali e não faltará quem diga que aquilo é obra de contestatário.

As Notas de Bensaúde transcendem a realidade Técnico e são sem tempo.

Há muitos anos emprestei a um amigo, mais idoso, culto mas que não conhecia o livro, o Conde de Abranhos... do Eça.

É espantoso como deixaram publicar! Embora escrito pelo Eça de Queiroz. É espantoso como é fácil ter-lhe acesso.

Assim, são essas notas de Bensaúde, plenas de sabedoria, de verdade, e em que se adivinham tempos difíceis de luta e sacrifício.

Detesto histórias que à guisa de conclusão rematem com a moral da dita história, respeito a Moral mas contento-me com os factos, acho até antipedagógico que se saliente a moral.

Leiam, leiam, as Notas Histórico-Pedagógicas de Bensaúde: alunos, professores. Todos aqueles que se preocupam com a Escola, com a Universidade, com o Ensino e Investigação.

E difundam, difundam por todos os outros que sem grande preocupação e menor ocupação também falam da mesma Universidade.

Em Bensaúde a vocação de professor era dominante, é um facto.

Mas a força de que dispôs para realizar a sua obra; essa foi encontrá-la Bensaúde na crença que tinha na importância que reconhecia ao acto de ensinar, ao professor.

E também na crença que, em abstracto, tinha nos alunos:

"Há quem explique os fracos resultados obtidos em algumas das nossas escolas, como consequência da má qualidade dos alunos portugueses; é cómoda explicação...".

"A experiência de muitos anos permite-me asseverar que o estudante português é, em média, tão bom como o dos diversos países mais cultos da Europa..."

Com honra, prazer e alguma utilidade passei pelo Instituto de Alta Cultura. Assim foi-me possível trazer até nós 9 estudantes japoneses que na Universidade de Sofia, em Tokio, seguiam a licenciatura em Filologia Portuguesa.

Foi agradável recebê-los. De facto foi inesquecível.

Falaram um bom português. Japonês para mim seria impossível. Havia o inglês. Mas em português, em português era o ideal.

Em vez de lhes falar de Portugal, perguntei-lhes pelo Japão.

Julgo que é um país onde respeitam o professor?

Entreolharam-se, uniram-se, falaram baixo e foi uma moça que me respondeu:

Sim, é verdade: nos campos, nos teares, no teatro, na escola, todo aquele que ensina o que sabe é respeitado, é o professor.

Ofereceram-me um livro, The Southern Barbarians, The First Europeans in Japan.

Éramos nós. Lá chegámos nos meados do século XVI e ainda hoje nos agradecem a modernidade que então fomos capazes de lhes levar. Apropriaram-se do nosso modo de navegar, dos métodos de traçado das nossas cartas.

O caso japonês é bem conhecido.

Já antes. Mas sobretudo, depois da Primeira Grande Guerra, a partir de 1919, o Japão investiu poderosamente no ensino e na investigação.

Visitou e enviou os seus filhos para centros de excelência do estrangeiro.

Foi visitado, fez-se visitado por numerosos cientistas.

Outra guerra, diversos alinhamentos políticos, até duas bombas atómicas lhes aconteceram.

Mas as Leis da Física são sempre as mesmas e hoje aí estão. Aí está o Japão e o que faz.

Ensinar é uma forma de dar. Se dar significa ficar com menos, nem sempre se dá.

Alguém, desta casa, me perguntou se eu ensinava tudo quanto sabia.

Tudo quanto souber em consciência e utilidade ensino, ensino tudo.

Olhe que isso é perigoso!

Lembrei-me, ocorreu-me então dizer que um bom professor assemelha-se, talvez, a um poço, dá de beber, esvazia-se um pouco, esvazia-se completamente, mas enche-se naturalmente, enche-se de novo e atinge o nível habitual no momento oportuno.

Diferente é um depósito. Sem fontes próprias, que se enche custosamente e ainda mais custosamente se torna a encher.

Mas pode haver outros motivos para não dar. Infelizmente há muitos.

Simples hipótese: uma ideia que sendo nossa se cede a outros.

Imaginem Pasteur negociando, retardando a aplicação das suas investigações sobre a vacina.

Talvez por isso, Pasteur dizia que não havia ciências puras e aplicadas, mas tão somente, aplicações da Ciência.

E, sendo assim, como essa Ciência tem que ser pura, livre! Facultada a tudo e a todos a quem possa servir.

Segredos de conveniência, competição nos lucros auferíveis, haverá sempre. Em todas as épocas, em todos os quadrantes.

Mas cuidado!

Deverá haver locais onde a generosidade, a isenção, a independência sejam o quotidiano e o prevalente. Na Universidade em que me vejo deve ser assim.

E para tanto não se define e ainda menos se quer uma Torre de Marfim.

Na realidade a Universidade não pode alhear-se da sociedade onde se insere pela simples razão de, em princípio, a Universidade ser onde se sabe mais.

Erro será não recorrer e não ajudar a Universidade. Grave será se, em qualquer medida, perder a sua independência.

A Universidade ensina o que aprende através de uma actualização e investigação permanentes e continuadas.

Mas o facto primeiro, aquele que deontologicamente lhe cabe, é ensinar.

Ensinar àqueles que frequentam as licenciaturas e os seus cursos de pós-graduação.

Ensinar a quem lhe pedir conselho.

Colaborar na pesquisa de novas soluções. Sempre ensinar.

No nosso entender o acto primeiro é ensinar e antes desse, anónima, difícil, desesperante por vezes, está a aprendizagem do que se vai ensinar, a investigação do facto novo com que se depara, o esclarecimento das contradições que se constroem.

As leis de Newton são do domínio público, as equações de Maxwell não combram direitos de autor.

O que torna legítimo o ensino destas matérias é tornar nosso o que foi descoberto e escrito por outros, é a assimilação séria e aprofundada do que se ensina.

Mas também cuidado!

Não confundamos essa ciência com a dispensa de se escreverem livros adequados.

É estranho, um pouco bizarro, que se usem como livros de um curso, livros contemporâneos escritos por outros, nossos pares.

Livro e professor devem constituir-se num todo harmonioso e em que cada parte é essencial.

Não se confunde isso com bibliografia recomendada que deve ser abundante e acessível.

E Bensaúde adverte:

"Além do conhecimento do que ensinam os compêndios, o professor, para produzir bons discípulos, precisa de outras aptidões, mais raras...

... Pasteur, que era um fraco orador, se tivesse nascido português, teria passado a vida, provavelmente, como figura secundária em algum dos nossos obscuros laboratórios. Num ambiente mais propício, conseguiu ser um grande sábio e um grande professor, embora a retórica não fosse o seu forte. 'Quando a sabedoria vai adiante a eloquência segue-a como fiel companheira' (Santo Agostinho)."

E sobre o aluno salienta:

"O nosso aluno normal possui uma rapidez de compreensão igual e, às vezes, talvez superior à dos rapazes que frequentaram as escolas da Europa Central. É esta a maior das suas qualidades inatas; as que deveriam ser desenvolvidas pela educação faltam-lhe, porém, muitas vezes. O seu poder de assimilação é quase sempre muito inferior ao dos estudantes cuja educação foi cuidada desde a infância. Pode dizer-se até que entre os nossos estudantes, a faculdade de assimilação costuma ser inversamente

proporcional à faculdade de compreensão; quanto mais facilmente compreendem, mais facilmente esquecem. É que compreender e fixar o que se compreende são coisas diversas."

Sem Brito Camacho, Alfredo Bensaúde não teria fundado o Técnico, a Escola Nova em que o Instituto se converteu.

Sem Duarte Pacheco, o Técnico nunca se teria desenvolvido como se desenvolveu. Isto, além do modo como Duarte Pacheco serviu a Engenharia Portuguesa e uma vez mais o Técnico.

Sem condições mínimas, os espíritos fogem da carne frágil que os não agarra.

Que se cuide do professor e da sua dignificação.

Que se cuide do aluno e da sua dignificação.

Deles depende o que for, mas outros, por dever de cargo, poderão e devem ajudar o que em muitos casos é de importância vital.

A Associação dos Estudantes do Instituto Superior Técnico faz 75 anos.

Criado por Decreto em 23 de Maio de 1911, o Técnico iniciou as aulas em 13 de Novembro. Seguiu-se a Associação dos Estudantes, fundada logo a seguir, 11 de Dezembro, exactamente como hoje.

Desde sempre a Associação e o Instituto seguiram caminhos paralelos ou de cruzamento próximo com um destino comum: a perenidade do I.S.T.

Pela Associação passaram alunos distintíssimos que ocuparam lugares na Direcção, na Técnica, nas muitas e diversas áreas em que a associação se diversificou e aplicou.

Passaram outros que não tendo sido alunos distintos foram, no entanto, hábeis gestores, impulsionadores e colaboradores da Associação.

Foi sempre característica da Associação dos Estudantes do I.S.T. o seu espírito fraterno de camaradagem, entre si e com outras associações académicas.

Foi sempre característica da Associação a sua inteira disponibilidade para com a Escola oferecendo-se em generosa colaboração.

De tal modo considero importante a acção das Associações de Estudantes que entendo por desnecessário defender a sua existência.

Além de tudo quanto realizou no campo pedagógico, cultural e desportivo. Além da realidade forte que é a TÉCNICA, além de tudo isso, a Associação foi e será sempre uma escola de dirigentes, não tenhamos dúvidas do facto, nem do que daí resultou e resultará.

Houve momentos difíceis, de diálogo impossível e a maior parte das vezes impossibilitado. Mas houve sempre retorno amigo à colaboração.

Política não é exactamente agitação nem outras coisas feias. Política, o Governo da Cidade, é uma coisa bela em que todos devem comparticipar para que não sejam governados como não querem.

É uma coisa bela quando praticada nos exactos preceitos seguidos por uma Universidade independente e pura.

As Associações foram e serão sempre origem natural de mentalidades robustas e críticas. De intervenção possível e desejável nos sistemas educativos.

Tenho a grande honra e acrescentado prazer de ser Director da Técnica.

Na Técnica se publicaram, muitas vezes, as novidades científicas que, então, ainda incipientes, ganhariam vulto e extraordinária dimensão e influência.

Essa consciência científica da Técnica deve-se aos professores do Instituto que nela colaboraram desde a primeira hora.

A Técnica é uma revista portuguesa, de estudantes, escrita por professores e não exactamente de modo total para os estudantes mais novos. É um bom investimento, no entanto, para o futuro e uso desses mesmos estudantes.

Numa tendência para certo derrotismo, numa vaidade sem ter de quê, nem sempre analisada até às raízes onde se alimenta, nem sempre se acorre à Técnica. É pena. É mau para o pensamento científico português. É mau exemplo que não ajuda a frutificar outras revistas similares.

No entanto, a Técnica viaja, pedem-nos separatas de além-fronteiras.

Termino.

Associação é uma palavra bela, plena de intenções.

Estudante, ser estudante, é belo e sem idade.

Os jovens terão sempre por si o futuro que postulam e invadem e para o que, em boa verdade, não há experiência que invalide os sonhos com que, no presente, o controem.

A vida continua-se com a chegada dos que nascem.

Estudante na escola é episódico e efémero.

Que a Escola e a sua Associação consigam formar estudantes de sempre e associados de sempre.

Isso significa a perenidade da trilogia:

Instituto Superior Técnico.

Associação dos Estudantes do I.S.T.

Associação dos Antigos Alunos do I.S.T.

Tão só isso!

Palavras proferidas na Sessão de encerramento das comemorações dos 75 anos da Associação de Estudantes do I.S.T. por

FRANCISCO SANCHES

da Direcção da AEIST

Celebrar 75 anos de vida de uma organização, é necessariamente, reflectir no passado, analisar o presente e projectar o futuro.

Quando se trata duma Associação fundada e gerida por jovens universitários, junta-se ainda a obrigação de realizar cada vez mais e melhor.

Em 11 de Dezembro de 1911 os objectivos estatutários da Associação de Estudantes do Técnico eram os seguintes:

"a) Difundir a instrução entre os associados, promovendo excursões e visitas de estudo, conferências literárias e científicas e especialmente as que se relacionam com o estado financeiro, industrial, comercial e agrícola do nosso país e tudo o mais que possa concorrer para o aperfeiçoamento moral e físico dos seus sócios.

b) Procurar obter regalias que tornem as condições dos associados mais fáceis."

Como se nota são objectivos que transportados para os nossos dias não perderam actualidade.

Pensamos que ao longo do tempo a Associação de Estudantes do IST tem sabido cumprir as ideias de solidariedade e de ajuda aos estudantes do Técnico.

Depois de uma fase inicial em que acompanhou uma certa indefinição da escola, em termos de fins e instalações, a Associação de Estudantes passou a dispor de pavilhão próprio, passo importante para a sua consolidação e expansão.

Seguiram-se no tempo fases opostas que "adormeceram" um pouco a acção da Associação. Referimo-nos ao controlo dos corpos gerentes exercido pelo Estado e ao fervor revolucionário após Abril 1974. Foram fases das quais é importante tirar os ensinamentos necessários para não cair nos mesmos erros.

Actualmente, a Associação de Estudantes do Técnico apresenta, uma estrutura dividida por várias áreas de intervenção: Associativa, Federativa, Desportiva e Pedagógica subdividida por secções amadoras que promovem e sustentam as actividades extracurricular dos alunos do I.S.T. Efectivamente, é no parque desportivo e nas suas restantes instalações que é dinamizada a actividade recreativa/cultural e desportiva dos estudantes. Infelizmente, as cargas horárias continuam a não permitir um acesso regular às actividades extracurriculares.

Se no plano interno se reconhece que é necessário melhorar os serviços prestados, é no plano externo que se consolida a posição da maior e mais importante Associação de Estudantes da Academia de Lisboa. Não é por acaso que as duas estruturas federativas mais recentes, Associação Académica de Lisboa e a Rádio Universidade Tejo são lideradas por representantes da nossa Associação.

Outro vector importante de actuação é o da intervenção pedagógica contribuindo para a melhoria do ensino no I.S.T. Pensamos que será aqui que a Associação deverá intervir no futuro ainda com mais empenho e maior participação de todos os estudantes.

Há que ultrapassar a velha luta aluno-professor avançando para a resolução dos problemas através de ideias e projectos que indiquem que a nossa função não é a da contestação pura e simples. Apostar num ensino melhor é a meta que a Associação de Estudantes do I.S.T. deve atingir o mais depressa possível. No entanto, isto só se consegue se houver a participação de todos, através de sugestões e críticas.

A Associação de Estudantes do Técnico será aquilo que todos os Estudantes desejarem, pois, há sempre a possibilidade de mudar os seus dirigentes e avançar com alternativas.

Continuar a obra iniciada em 11 de Dezembro de 1911 não é tarefa fácil. Mas se todos quiserem teremos uma Associação melhor e mais participada.

Palavras proferidas na Sessão de encerramento das comemorações dos 75 anos da Associação de Estudantes do I.S.T. por

VENÂNCIO ROSA

da Comissão Organizadora das Comemorações

O motivo que nos trouxe aqui tem algo de especial, pois a Associação de Estudantes do I.S.T. faz hoje 75 anos.

Desde 11 de Dezembro de 1911 — data da sua fundação — até aos nossos dias esta Associação de Estudantes tem sido um marco na vida académica do País, nunca deixando que a sua voz esmoreça ou seja abafada sempre que estão em causa os direitos que a todos os estudantes devem assistir, e especialmente na cidade de Lisboa onde, pelo seu dinamismo muito tem contribuído para estreitar os laços que unem as diversas faculdades onde temos, como exemplo recente, o facto de ser uma das Associações de Estudantes que impulsionou a formação da A.A.L.

Por tudo isto e porque são poucos os organismos deste tipo que se conseguem manter durante tantos anos com tão forte personalidade, não quisemos deixar de comemorar convenientemente estas Bodas de Diamante.

Estas comemorações tiveram como principais objectivos:

— mostrar o que os alunos fazem, hoje, em cada curso, no Técnico e em que termos a indústria e o País podem contar com os futuros engenheiros;
— abrir as portas para o exterior, especialmente para a cidade;
— promover o convívio entre universitários das diversas faculdades que, como todos sabemos, é bastante deficiente, devido a estas estarem muito dispersas — facto que muito lamentamos.

Para conseguirmos estes objectivos realizámos um vasto leque de actividades que tiveram início no dia 17 de Novembro e se prolongaram até hoje, das quais passo a fazer uma breve retrospectiva.

De 17 a 22 de Novembro decorreu a exposição "A Engenharia no I.S.T." que, contando com equipamento de laboratório dos vários cursos e com muitos trabalhos de final de curso executados pelos alunos mostrou a todos os visitantes, que foram muitos, as grandes potencialidades que o engenheiro licenciado no Técnico tem, tanto para a indústria como para a investigação.

De 22 a 30 de Novembro realizaram-se diversas actividades das quais passo a referir algumas:

— Torneios inter-faculdades de basquetebol, voleibol, futebol-5, andebol, pólo aquático, uma prova de atletismo e mesmo um torneio de xadrez, o que movimentou centenas de universitários.
— Cinema — já é tradição passarem-se bons ciclos de cinema no I.S.T. Esta foi mais uma ocasião para a confirmação deste facto, contando-se com três filmes que tiveram uma afluência de público além das espectativas.
— Teatro — não sendo muito comum o gosto pelo teatro no I.S.T. o professor Listopad e o Grupo de Teatro da U.T.L. vieram mostrar-nos como se faz teatro universitário.
— Noite de Coimbra — é quase um hábito contarmos nas nossas festas e comemorações com os colegas de Coimbra cuja tradição marca a vida académica do País. Contámos nessa noite com a presença da Tuna, Órfeon e um grupo de fados e cantares da Associação Académica de Coimbra que nos deliciaram com a boa música e cantares de Coimbra e contou com um público caloroso que esgotou a lotação da sala.
Melhor ainda foi o Convívio que se estabeleceu entre todos os elementos, cerca de 100 e os alunos do Técnico que com eles contactaram.

— Baile do Engenheiro — com este baile quisemos reatar a tradição dos bailes de gala nesta casa, mas teve como objectivo principal o estreitar das relações entre estudantes, professores e antigos alunos.

No dia 6 de Dezembro realizámos um espectáculo na Aula Magna, pois, não quisemos deixar passar estas comemorações sem fazer referência à música popular portuguesa, por isso, organizámos um espectáculo com o cantor Fausto que cantou com a sala cheia, verificámos com alegria, que a maioria eram estudantes.

A Comissão Organizadora destas comemorações e a D.A.E.I.S.T. querem agradecer a todas as pessoas que ajudaram na organização e colaboraram connosco e, muito especialmente, aos membros da Comissão de Honra, que sem o seu apoio e avalo não tinha sido possível a realização de todas estas actividades.

Queremos também aproveitar esta oportunidade para prestar homenagem a todos os dirigentes associativos que ao longo destes 75 anos trabalharam para que a nossa Associação de Estudantes seja o que é hoje e, principalmente, aqueles que não querendo deixar que o lema destes 75 anos "Construir um futuro" se ficasse pela sua vida académica, se distinguiriam, por isso, na vida pública nacional.

FÍSICA DE ALTAS ENERGIAS: Actividade recente do Grupo de Física Experimental de Altas Energias de Lisboa

José Mariano GAGO*, Amélia MAIO, Mário PIMENTA*, João VARELA***

RESUMO

Apresenta-se sucintamente algumas das experiências de física de altas energias no CERN nas quais participa o Grupo de Física Experimental de Altas Energias de Lisboa, indicando, para cada uma delas, os objectivos propostos, o equipamento utilizado, e os resultados alcançados.

Apresenta-se, igualmente, um trabalho de desenvolvimento tecnológico recente no domínio das fibras ópticas cintilantes, motivado pela necessidade de novos detectores em física de altas energias.

1. INTRODUÇÃO

O Grupo de Física Experimental de Altas Energias de Lisboa é um colectivo de trabalho formado por investigadores de várias instituições para actividades de pesquisa no domínio da física das partículas elementares. O Grupo realiza parte do seu trabalho em Lisboa, nomeadamente a concepção de experiências, desenvolvimento tecnológico, desenvolvimento de "software" e análise de resultados, executando no CERN o que diz respeito à preparação directa das experiências e à tomada de dados, em colaboração com grupos de outros países[1].

Este artigo visa apresentar algumas das actividades recentes do Grupo.

Falaremos assim de três grandes tópicos. Em primeiro lugar, focaremos uma experiência em preparação, cujo período de tomada de dados terá lugar em finais de 1986 e início de 1987, destinada à investigação da possível energia. Esta experiência, denominada NA38, detectará pares de muões emitidos nessas colisões e medirá ainda a energia neutra emitida. Sucede à experiência NA10, na qual também participam físicos portugueses, onde foram feitas descobertas importantes sobre o comportamento dos quarks e gluões no interior da matéria hadrónica.

Em segundo lugar, descreveremos uma experiência já realizada e concluída, destinada a esclarecer a natureza do pico na secção eficaz de retrodifusão das colisões antiprotão-protão, na região dos 180°, aparentemente idêntico ao crescimento que se observa em óptica nas secções eficazes de retrodifusão da luz pelas gotas de água e que é responsável pelo efeito chamado de "auréola". A experiência permitiu mostrar que esse pico da secção eficaz antiprotão-protão não é de facto explicável como efeito de auréola de tipo óptico; permitiu ainda chamar a atenção para um comportamento em energia das secções eficazes de certos canais, cujo esclarecimento exigirá outros dados experimentais.

A mesma experiência WA74 conduziu a um resultado inesperado em domínio diverso daquele para que fora inicialmente projectada. Com efeito, foi possível estudar, com precisão nunca atingida anteriormente, a polarização dos lambdas produzidos na região extrema para a frente (x 1) com feixe incidente de K^-. A polarização medida (que é, em última análise, a do quarks difundido a pequeno ângulo na matéria hadrónica) cresce fortemente com o momento transverso do lambda produzido. A razão deste importante fenómeno (só agora descoberto) não é conhecido.

Em terceiro e último lugar, descrevemos o trabalho

* Instituto Superior Técnico (Departamento de Física) e CFMC (INIC).
** Faculdade de Ciências de Lisboa (Departamento de Física) e CFNUL (INIC).

[1] A composição do Grupo de Física Experimental de Altas Energias de Lisboa é a seguinte (Novembro 1985): a) Investigadores: Conceição Abreu (FCL), Gaspar Barreira (CFNUL), Paula Bordalo (IST), Augusto Casaca (IST-INESC), José Mariano Gago (IST), Amélia Maio (FCL), Mário Pimenta (IST), Sérgio Ramos (IST), João Varela (IST); b) Estagiários: Carlos Cruz (IST), Paulo Filipe (IST), Luís Peralta (FCL), Pedro Vaz (IST). O GFEAEL é coordenado por José Mariano Gago e funciona em Lisboa, no Complexo II do INIC (Instituto de Física e Matemática), Av. Prof. Gama Pinto, 2, 1699 Lisboa Codex, Tel.: 77 33 25/77 33 38/77 42 97, Telex 62593 IIFM P.

Original recebido para publicação em 20/11/85.

desenvolvido no domínio da utilização e do estudo de fibras plásticas cintilantes, cujo sucesso parece assegurado para a realização de certo tipo de detectores, nomeadamente de calorímetros electromagnéticos e de hodoscópios de cintilação de grande resolução.

2. A EXPERIÊNCIA NA38 (em preparação): PROCURA DE PLASMA DE QUARKS E GLUÕES[2]

A experiência NA10 sondou a estrutura da matéria hadrónica estudando as colisões de piões de 200 GeV sobre um alvo de tungsténio onde são produzidos pares de muões de carga contrária. Estes são coerentemente interpretados como o resultado da aniquilação de um quark com um antiquark, respectivamente de um nucleão no alvo e de um pião incidente, num fotão virtual massivo que se materializa num par de muões. Os resultados confirmam que os hadrões são partículas com estrutura, dão a distribuição de impulsão dos quarks no seio dos piões e nucleões e põem em evidência alguns dos aspectos da dinâmica da interacção forte.

Esta visão da matéria composta de quarks estruturou-se de facto com base nos resultados de numerosas experiências diferentes, embora nenhuma das tentativas realizadas para observar os quarks no estado livre tivesse resultado. Actualmente julga-se compreender este facto como o resultado da natureza da interacção forte que obrigaria os quarks a manterem-se confinados no interior dos hadrões.

A matéria pode, no entanto, organizar-se diferentemente quando submetida a condições extremas de temperaturas e pressão, tais como as que existiram nos primeiros instantes da criação do Universo (A Grande Explosão) e as que existem actualmente no interior das estrelas de neutrões. De facto, a teoria que os físicos dispõem para descrever a interacção entre os hadrões (Cromodinâmica Quântica), quando extrapolada para densidades de energia superiores à observada na matéria nuclear ($\epsilon_N = 0.15$ GeV/fm^3) prevê a existência de uma transição de fase para um outro estado da matéria, o plasma de quarks e guões desconfinados. Quando a densidade de energia atinge valores elevados, quer devido a uma compressão extrema, caso das estrelas de neutrões, quer por via de grandes energias de agitação térmica (temperatura), caso dos instantes iniciais do Universo, o campo de cor (carga da interacção forte) sofre um efeito de blindagem "tipo Debye", de forma que os hadrões perdem a sua identidade e os quarks e guões coloridos passam a formar um gás de partículas quase-livres. As estimativas existentes indicam que a transição de fase se dá para uma densidade de energia da ordem de 2 GeV/fm^3, o que no plano temperatura-densidade de número bariónico corresponde à linha esboçada na fig. II.1.

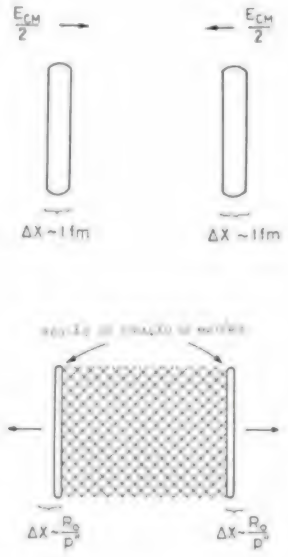


Fig. II.1 a) Representação esquemática de dois núcleos observados no referencial do Centro de Massa antes da colisão. As dimensões longitudinais são da ordem de 1 fm; b) Região de formação de matéria após a colisão.

Num futuro próximo abre-se a possibilidade de reproduzir em laboratório as condições iniciais do Universo, embora, como é óbvio à escala das dimensões nucleares. De facto, preparam-se actualmente, no CERN experiências destinadas ao estudo das colisões com iões de oxigénio acelerados a 200 GeV/nucleão no SPS (Super Proton Syncrotron). Nas colisões mais violentas, estima-se que sejam produzidas da ordem de 400 partículas e que se atinjam densidades de energia da ordem de 2 GeV/fm^3. Para compreender o que se passa quando dois nucleões colidem a alta energia é necessário fazer apelo à teoria da relatividade. Observados no referen-

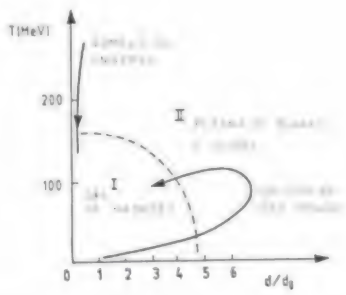


Fig. II.2 Representação da frase hadrónica e do plasma de quarks e gluões no diagrama temperatura-densidade de número bariónico.

[2] A colaboração científica responsável pela experiência NA38 é a seguinte: Bergen, CERN, Clermont-Ferrand, Ecole Polythechnique, Lisboa (INIC), Lyon, Neuchâtel, Orsay, Strasbourg, Valência.

Os signatários portugueses desta proposta de experiência são: C. Abreu, G. Barreira, P. Bordalo, A. Casaca, J. M. Gago, A. Maio, M. Pimenta, S. Ramos e J. Varela.

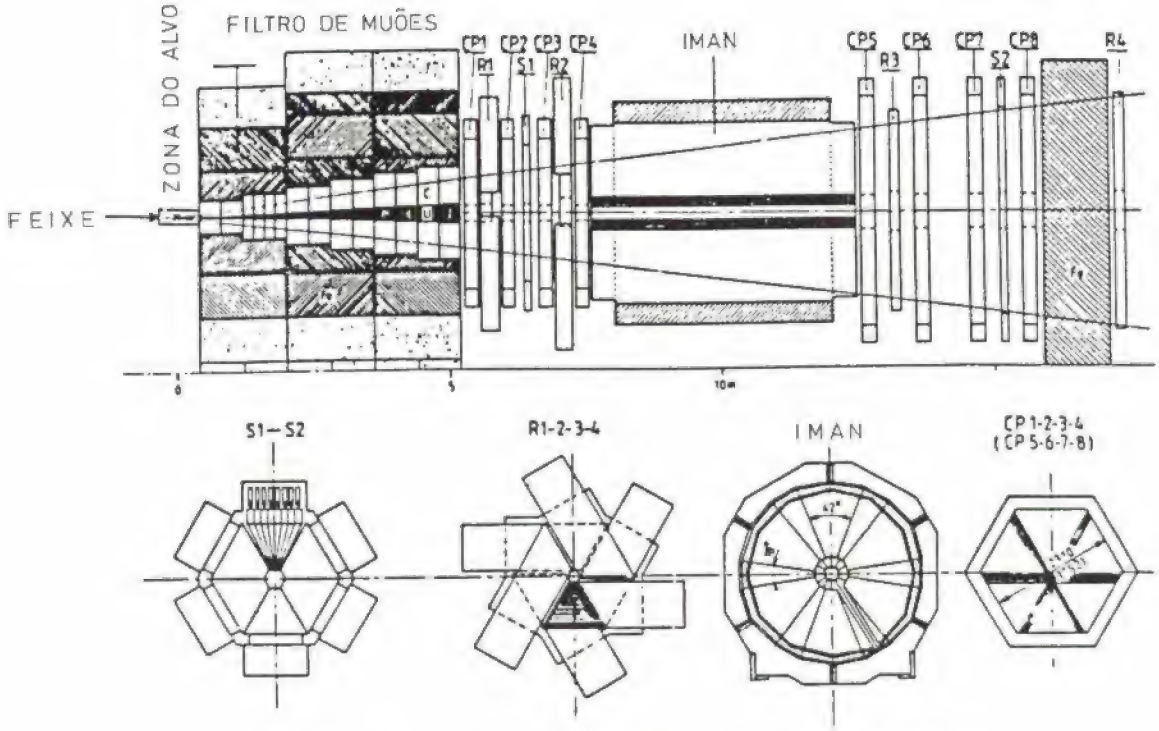


Fig. II.3 Espectrómetro de muões da experiência NA.10.

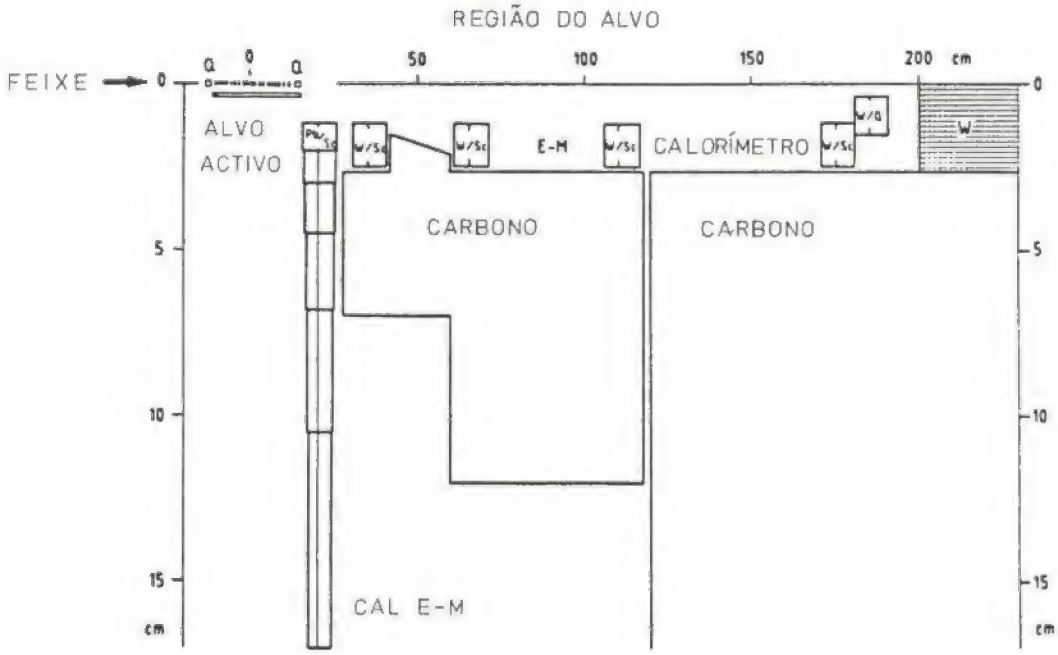


Fig. II.4 Detectores na região do alvo.

cial do centro de massa da interacção, os dois núcleos, por contracção relativista, apresentam a forma de duas bolachas movendo-se em direcções opostas (fig. II.2). As dimensões longitudinais dos núcleos, que no limite da velocidade da luz se anulariam são limitados a 1 fm pelo princípio de incerteza. O tempo de colisão, é portanto, da ordem de 1 fm/c. Nas colisões são criadas partículas num tempo próprio da ordem das suas dimensões, 1 fm/c. No referencial do centro de massa este tempo é dilatado relativisticamente resultando, portanto, consideravelmente superior ao tempo de colisão. Desta forma os núcleos são transparentes no momento de interacção, ocorrendo a criação de matéria (quarks e gluões) em instantes posteriores, na região delimitada pelos núcleos em afastamento. Se a densidade de energia nesta região for inferior à densidade crítica, os quanta de matéria organizam-se en hadrões constituindo um gás de partículas que se expandirá até ser detectado pelo equipamento experimental. No caso contrário não há lugar para a imediata condensação hadrónica, de forma que os quarks e gluões permanecem livres formando um gás sujeito a dois processos em competição, um de termalização conducente à uniformização da temperatura, e outro de expansão, essencialmente longitudinal, governado pelas equações da hidrodinâmica relativista. Em tempos posteriores do processo de expansão o gás arrefece, a densidade de energia desde abaixo de ponto crítico, e os quarks recombinam-se dando origem aos produtos finais da reacção.

Apercebemo-nos facilmente de que a complexidade do processo de colisão de dois núcleos relativistas obscurece a eventual produção do plasma de quarks e gluões, tornando delicada a sua identificação e caracterização. Com efeito, a grande maioria dos produtos da reacção reflecte, essencialmente, a física que ocorre na frente de hadronização ao longo da expansão do plasma. No entanto, o gás de quarks e antiquarks deve ser a sede de abundante produção de fotões, distribuídos segundo a lei de Planck se aquele se encontrar termalizado. Estes fotões, tanto reais como virtuais, são um verdadeiro termómetro indicador das características do plasma e da natureza da transição de fase.

Do ponto de vista experimental, a detecção de fotões reais provenientes do plasma é seriamente dificultada pela presença dos fotões de desintegração dos piões neutros, abundantemente produzidos no processo de harmonização. Em contrapartida, os fotões virtuais materializam-se em pares de muões de fácil detecção e medida. A experiência NA38 no CERN visa, assim, a medida dos pares de muões de massa superior a 0.5 GeV/c^2 produzidos nas colisões oxigénio-urânio, acompanhada da medida da distribuição de energia hadrónica por forma a determinar o grau de inelasticidade das colisões.

A detecção de pares de muões é feita pelo espectrómetro da experiência de NA10 (fig. II.3). O princípio de detecção baseia-se na diferença das probabilidades de interacção com a matéria dos muões e dos hadrões. Um bloco de carbono de 4,8 m após o alvo absorve a totalidade dos hadrões produzidos conjuntamente com os muões, que são desta forma isolados e medidas pelos detectores colocados a jusante. A medida da impulsão é feita por deflecção das trajectórias num campo magnético toroidal, as quais são reconstruídas a partir da informação recolhida em câmaras de fios. Os acontecimentos interessantes são sinalizados por lógica electrónica de grande rapidez, com base nos sinais fornecidos por um conjunto de 800 cintiladores reunidos em 4 hodoscópios.

A medida dos hadrões é realizada por um detector compacto, instalado entre o alvo e o filtro de muões, especialmente concebido para esta experiência (fig. II.4). Um calorímetro electromagnético mede a distribuição de energia dos fotões produzidos, os quais resultando essencialmente do decaimento $\pi^0 \rightarrow \gamma + \gamma$ nos dão uma medida dos piões neutros. Os 36 módulos do calorímetro são constituídos por camadas sucessivas de chumbo e fibras ópticas cintilantes de 0.5 mm de espessura. A luz recolhida nas fibras cintilantes é proporcional à energia dos fotões incidentes iniciadores de cascatas electromagnéticas no interior do calorímetro. Os sinais luminosos são conduzidos por guias de luz até fotomultiplicadores onde são convertidos em impulsos eléctricos, posteriormente digitalizados por conversores analógico-digitais.

O alvo de urânio subdividido em 20 segmentos de 1 mm de espessura espaçados de 1 cm, é envolvido por 25 cintiladores anelares de 0.7 cm de diâmetro. A altura de impulso medida nos cintiladores é proporcional ao número de partículas carregadas que os atravessam. Este sistema cumpre duas funções, por um lado mede a distribuição angular das partículas carregadas e por outro identifica eventuais reinteracções de fragmentos nucleares.

O feixe incidente, cuja intensidade é de 10^8 iões/seg, contém uma pequena proporção (1%) de iões de cálcio, para além de fragmentos de menor número atómico (Be, alfa...). É, portanto, necessário dispôr de um sistema capaz de medir intensidades elevadas e de identificar os iões incidentes. Estas funções são realizadas complementarmente por um hodoscópio de 28 cintiladores colocado no percurso do feixe e por dois detectores de quartzo colocados à entrada e à saída do alvo.

O conjunto dos sinais analógicos fornecido pelo equipamento adicional (calorímetro, detector de alvo, detector de feixe) é analisado em amplitude por electrónica apropriada e transferido, conjuntamente com os sinais lógicos provenientes do detector de muões, para um sistema de computadores Nord 100/500, onde são registados em banda magnética para análise posterior.

O Grupo de Física Experimental de Altas Energias de Lisboa, (GFEAEL) participou activamente na elaboração da proposta de experiência NA38, recentemente

aprovada pelo Comité de Pesquisa do CERN, nomeadamente os estudos de adaptação do espectrómetro de muões às novas condições experimentais. Durante o próximo ano, a GFEAEL é responsável pelo desenvolvimento dos programas de aquisição de dados em linha, bem como dos programas de simulação e por testes do calorímetro electromagnético. O GFEAEL é também responsável, em colaboração com o INESC, pela construção de uma interface destinada a compatibilizar os dois sistemas de aquisição utilizados (RMH-ROMULUS).

3. A EXPERIÊNCIA WA74 (concluída)[3]

A experiência WA74 realizou-se entre Fevereiro e Maio de 1982 no espectrómetro Omega na zona oeste do CERN, urilizando um feixe não separado (e^-, π^-, k^- p^-) e um alvo de hidrogénio líquido (fig. III.1.).

O espectrómetro Omega dispõe de um íman supracondutor com um campo magnético vertical (~ 1.8T) que permite boas aceitâncias geométricas.

As trajectórias das partículas carregadas foram reconstruídas, na zoná para a frente, por um sistema de quinze câmaras proporcionais e duas câmaras de deriva e, na região de grandes ângulos por dezasseis câmaras proporcionais colocadas paralelamente ao alvo.

As partículas rápidas difundidas para a frente atravessam um ou dois Cerenkovs de limiar, que permitem estabelecer, para determinados intervalos em momento, hipóteses de massas, e dois ou três hodoscópios verticais.

O trigger da experiência exigia a existência de uma partícula rápida positiva de impulsão próxima da dos feixes e foi elaborado com a ajuda da informação recolhida por um sistema de leitura rápida instalada num plano duma das câmaras proporcionais.

Este trigger permitiu seleccionar candidatos às reacções a dois corpos $\pi^- p \to p\pi^-$, $pp \to \pi^+ \pi^-$, $k^- p \to pk^-$ e $pp \to pp$ e ainda estudar produção inclusiva de lambdas.

Nos parágrafos seguintes apresentam-se os principais resultados obtidos nestes dois tipos de reacções:

a) Estudos de reacções a dois corpos

Na reacção $\pi^- p \to p\pi^-$ a secção eficaz diferencial $d\sigma/du$ foi determinada entre $0<u<1\ GeV^2/c^2$. Os valores encontrados nesta recção enquadram-se nos valores publicados para momentos incidentes entre 6 e 90 GeV/c que são globalmente explicados num modelo de Regge pela troca de uma única trajectória com $\alpha = -(0.06 \pm 0.05) + (0.50 \pm 0.26)\,u$.

O valor encontrado para a secção eficaz na reacção $pp \to \pi^+ \pi^-$ a 12 GeV/c permitiu, comparando com resultados publicados a 5 e 6 GeV/c, estabelecer uma de-

[3] A colaboração responsável pela experiência WA74 é a seguinte: CERN, Universidade de Paris VI-LPNHE, Collège de France, Lisboa (INIC).

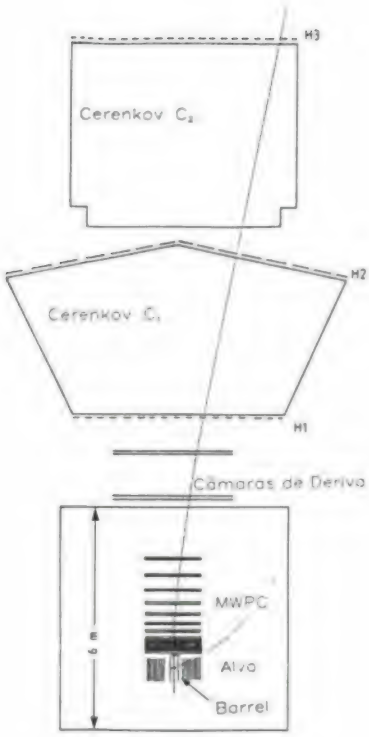


Fig. III.1 Dispositivo Experimental experiência WA74.

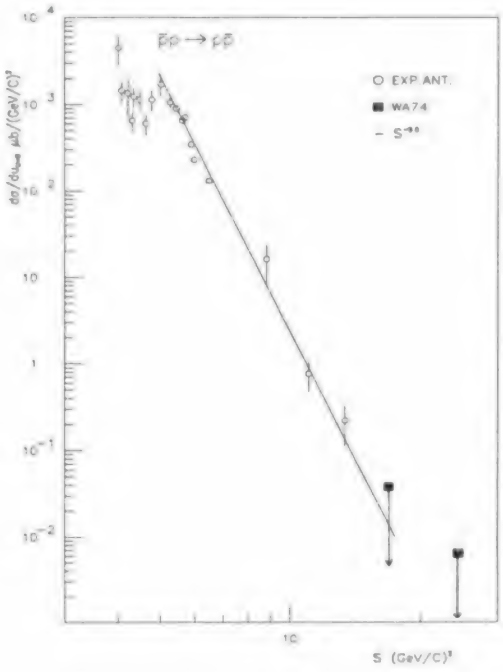


Fig. III.2 Secção eficaz diferencial na região para trás (u = O) em função da energia no centro de massa na reacção pp→pp.

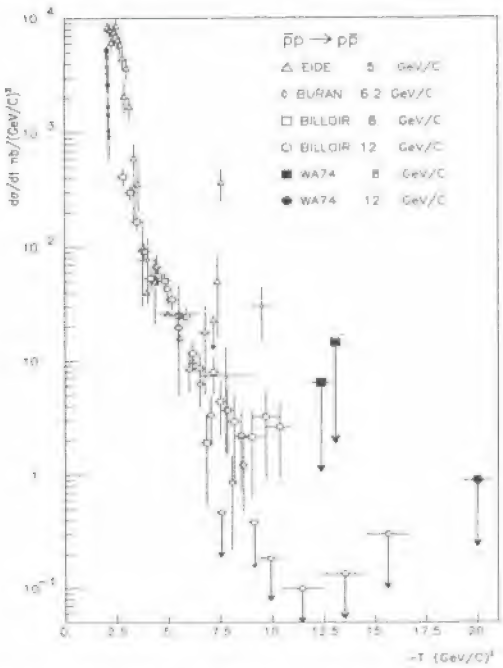


Fig. III.3 Secção eficaz diferencial dΘ/dt para vários momentos incidentges na reacção pp→pp.

pendência em energia análoga, dentro dos erros, à verificada na sua reacção cruzada ($\pi^- p \rightarrow p\pi^-$). Este resultado comprova pela primeira vez, para estas reacções e numa gama de energias significativa, uma previsão geral dos modelos de Regge pela qual as trajectórias trocadas em reacções cruzadas são idênticas.

Os limites superiores encontrados a 8 e 12 GeV/c na secção eficaz para trás na reacção $\bar{p}p \rightarrow p\bar{p}$ indicam a manutenção do regime de forte dependência em energia estabelecido anteriormente nesta reacção para momentos incidentes inferiores a 6.2 GeV/c (fig. III.2). De igual modo estes limites são, quando comparados com os resultados encontrados na mesma reacção e energias na região de grandes ângulos, compatíveis com o crescimento da secção eficaz diferencial $d\sigma/du$ para $u \simeq 0$ GeV/c^2 como tinha sido verificado a 5 e 6.2 GeV/c (fig. III.3).

Na reacção $k^- p \rightarrow pk^-$ os valores de sensitividade alcançados com o feixe $\bar{k}$ não permitem que os limites superiores obtidos na secção eficaz para trás tenham significado apreciável.

A coexistência, em reacções exóticas, de uma forte dependência em energia da secção eficaz para trás com o crescimento, na mesma região, da secção eficaz diferencial $d\sigma/du$ para $u \simeq 0$ Gev/c^2 corresponde a uma mudança de regime na dependência em energia na secção eficaz para trás passando-se de uma forte dependência para uma dependência suave. Este comportamento é associado, em modelos de Regge, ao aparecimento de uma trajectória dominante. Note-se, porém que, mesmo para os canais não exóticos, não existe qualquer modelo teórico que explique a zona de forte dependência em energia. Esta dependência é aliás análoga à encontrada nos canais exóticos na região para trás bem como nas reacções exclusivas na região de grandes ângulos.

O estudo exploratório do comportamento das secções eficazes na região para trás em reacções exóticas a energias mais elevadas foi efectuado nesta experiência. O problema físico subjacente, compreensão dos mecanismos dominantes nas zonas de forte dependência em energia, parece-nos interessante e merecedor dum estudo sistemático. Para tanto é indispensável um feixe separado e intenso de p e/ou k^- com um momento incidente entre 10 e 20 GeV/c que permita sensitividades de ~ 1 acontecimento por picobarn. De momento, não se encontram disponíveis feixes com estas características. Contudo, a existência de alguns projectos em estado de elaboração muito diferente (nova geração de acumuladores p, TRIUMF, HIPS, SUPER-LEAR...), permitem ter esperança que num horizonte de dez anos se possa modificar significativamente este panorama.

b) Polarização dos lambdas emitidos na região de fragmentação do K^-

Nesta componente da experiência WA74 a identificação da partícula do feixe foi obtida usando dois detectores Cerenkov na linha de feixe: um detector de limiar e um detector diferencial (CEDAR) ajustado à identificação de K^-. O "trigger" que selecciona um protão rápido emergente da colisão aceita lambdas no modo de desintegração $\Lambda \rightarrow p\pi^-$. A aceitância dessa selecção cresce de 10% a x = 0.6 até 70% a x = 1.0, onde x é definido como $x = P_l^*/P_{l_{max}}^*$, razão entre o momento longintudinal do lambda no centro de massa da colisão K^- p e o momento longintudinal máximo possível nesse referencial.

Após reconstrução e selecção, foram obtidos 16 630 lambdas a 12 GeV/c, mais de 90% dos quais com x>0.6. Apenas estes foram, finalmente, utilizados na análise.

A polarização dos lambdas pode ser estimada por $\langle \alpha P \rangle = \langle \cos\Theta \rangle / \langle \cos^2\Theta \rangle$, onde $\alpha = 0.642$ e $\cos = (K^- \times \Lambda) \cdot p_{cm\Lambda} / |(K^- \times \Lambda)| \cdot |p_{cm\Lambda}|$. Nesta fórmula, $p_{cm\Lambda}$ é o momento do protão de desintegração no referencial próprio do lambda.

Após uma análise exaustiva destinada a corrigir e minizar todos os possíveis erros sistemáticos introduzidos pela aparelhagem e pelos critérios e processos de reconstrução e selecção, obtiveram-se, no que diz respeito à polarização dos lambdas, os resultados que se mostram nas figuras III.4 e III.5.

Resumidamente, esses resultados mostram que os lambdas produzidos na região de fragmentação do K^- são polarizados na direcção $K^- \times \Lambda$. Essa polarização aumenta com o momento transverso do lambda (p_t) e é insensível a x na região considerada. A dependência da polarização com p_t pode ser ajustada à fórmula

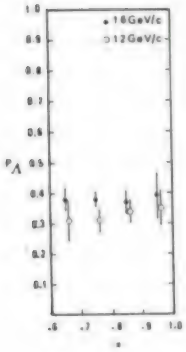


Fig. III.4 Polarização do Lambda em função de $x = Pl^*/P^*_{lmax}$ a 12 e 16 GeV/c de momento incidente.

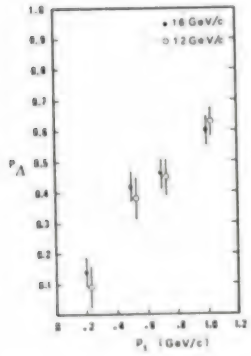


Fig. III.5 Polarização do Lambda em função do momento transverso do Lambda a 12 e 16 GeV/c de momento incidente.

$P(p_t) = Bp_t$ com o resultado $B = 0.66 \pm 0.03$ ($\chi^2 = 6.0/ND = 3$).

A importância deste resultado, que é inteiramente novo, reside no facto de a polarização dos lambdas na região de fragmentação do K^- ser, em última análise, a polarização adquirida pelo quark s, inicialmente não polarizado, ao difundir-se (e recombinar-se) na matéria hadrónica.

Um processo de polarização por difusão é conhecida, por exemplo, em electrodinâmica quântica, aquando da difusão de electrões não polarizados por um potencial coulombiano ou por um núcleo; a polarização é um resultado, nesse caso, dos diagramas de 2.ª ordem de QED (polarização de Mott). Podia, pois, legitimamente, perguntar-se se comportamento idêntico se não previria em cromodinâmica quântica. A resposta é negativa, ao menos no que diz respeito aos diagramas de 2.ª ordem: a dependência angular da polarização e a dependência em energia, nesse caso, são incompatíveis com os nossos resultados experimentais que permanecem, assim, e de momento, sem explicação.

4. FIBRAS ÓPTICAS CINTILANTES DE PLÁSTICO (desenvolvimento técnico para detectores de partículas)

As fibras ópticas cintilantes de plástico têm sido desenvolvidas para detecção de radiação electromagnética e em particular para calometria electromagnética em Física das Altas Energias. Por essa razão, o estudo das suas propriedades ópticas e de cintilação, e a sua resistência à radiação (U.V., γ, n, e^-) é da maior importância.

As fibras ópticas cintilantes[4] (f.o.e.p.) em estudo, são constituídas por um corpo central cilíndrico de poliestireno revestido de uma fina camada de acetato de polivinilo. O poliestireno é dopado com 1% de Butil-PBD, a componente cintilante, e com 10^{-4} de Demetil-Popop, o 2.º soluto, cuja função é deslocar a emissão no U.V. próximo para o visível. O D-Popop absorve a luz de comprimento de onda $\lambda \sim 375$ mm e reedite no azul de c.d.o. $\lambda \sim 430$ mm (fibra azul). Como o 2.º soluto orgânico outros compostos estão sendo ensaiados, como

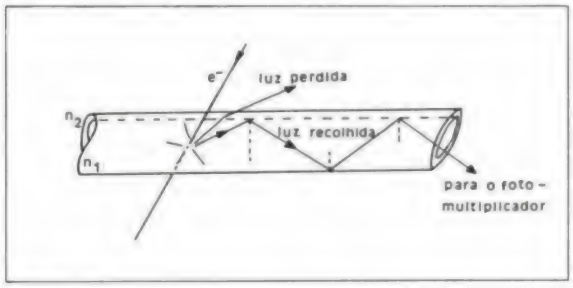


Fig. IV.1 Esquema da produção de luz e sua transmissão ao longo da fibra.

$n_1 = 1{,}59$; $n_+ = 1{,}46$; AN = 0,62.

sejam o BBOT, o BDB e o K27. Em particular o K27 que emite no verde (fibra verde) têm-se apresentado como possível alternativa ao D-Popop.

O esquema de produção de luz e sua transmissão devida ao fenómeno óptico de reflexão total, estão representadas na fig. IV.1.

Como se indica na fig. 1, n_1 é o índice de refracção do corpo central e n_2 é o índice de refracção do revestimento. Para estas fibras os valores dos índices de refracção são $n_1 = 1.59$ e $n_2 = 1.46$ e o valor da abertura numérica é de AN = 0.62. No momento actual são fabricados fibras com diâmetros que vão de 0.5 mm a 2 mm e a espessura do revestimento correspondente varia de 5 μm a 20 μm.

O objectivo deste projecto é pois, o estudo

i) das propriedades de cintilações das fibras ópticas de plástico cintilante, i.e., o rendimento na produção de luz quando a fibra é atravessada por radiação electromagnética, e

ii) das características de transmissão dessa luz através da fibra, medindo o comprimento de atenuação L_{at} da fibra, i.e., o comprimento de fibra

[4] Desenvolvidas em DP_hPE — STIPE, SACLAY, FRANÇA. Em Lisboa o trabalho aqui referido foi realizado por Carlos Cruz (CFMC), Amélia Maio (CFNUL) e Luís Peralta (CFNUL), sob a direcção da Prof.ª Amélia Maio.

atravessado pela luz tal que o número de fotões produzido inicialmente (n_o) é reduzido de

$$\frac{1}{e}\ (n(L_{at}) = \frac{n_o}{e}).$$

Esse estudo é feito para fibras de diferentes diâmetros e para diferentes dopantes. É do nosso conhecimento que as fibras se degradam sob o efeito de elevadas doses de radiação nuclear, e dado que nas situações físicas em que vão ser utilizados essas doses serão elevadas (~ 25 krad/h), é do maior interesse o estudo detalhado da *resistência destas fibras à radiação nuclear* (γ, n, e^-) medindo o rendimento fotónico e o comprimento de atenuação antes e depois da irradicação das fibras.

Por razões instrumentais é, em geral, necessário introduzir um guia de luz (material inactivo do ponto de vista de cintilação-plexiglas, por exemplo) entre a fibra cintilante e o fotodetector. Algumas técnicas foram ensaiadas para resolver este problema. Em colaboração com a Universidade de Neuchâtel (Suíça) e o CERN, estuda-se um novo método que consiste em irradiar parte da fibra activa com radiação U.V., destruindo a cintilação nessa zona. Deste modo a zona irradiada poderá funcionar como guia da luz produzida na zona da fibra não irradiada.

Aparelhagem experimental e resultados

A aparelhagem utilizada é típica em experiências de Física Nuclear.

Como fonte de electrões (β^-) utiliza-se uma fonte radioactiva de Sr^{90}. A luz produzida pela passagem dos electrões que atravessam a fibra óptica de plástica é colectada na janela de um fotomultiplicador (PM2) de baixo ruído e boa resolução. A distância da fonte radioactiva ao PM2 é variável permitindo estudar a transmissão da luz produzida através da fibra. A saída do PM2 está colocada entre duas outras fibras F1 e F3 de igual diâmetro e que juntamente com a electrónica associada servem de telescópio. Só são registados impulsos no Multicanal quando há "simultaneamente" (numa janela de 20 ns) detecção de luz nas fibras 1 e 3, i.e., o sinal de coincidência entre a via 1 e a via 3 é utilizado como sinal de disparo na porta do multicanal.

As propriedades das fibras são caracterizadas pelo número médio de fotoelectrões, fotões produzidos na fibra, ao ser atravessada por um electrão, transmitidos à distância *l*, e convertidos em electrões no fotocátodo. O número médio de fotoelectrões n_{fe} é medido em função da distância *l* e o seu valor obtém-se a partir da análise do espectro registado no multicanal no MCA é proporcional ao número de fotoelectrões produzidos em cada impulso).

O espectro de fotoelectrões que se observa no multicanal apresenta a forma gaussiana quando o número médio de fotoelectrões n_{fe} é grande. No entanto, para fibras de diâmetro $\phi \leqslant 1$ mm, esse número é inferior a 10 e daí os espectros apresentarem máximos e mínimos de distribuição mais complexa. De facto os espectros observados ajustam-se à convolução de uma distribuição de Poisson cuja média é n_{fe} com uma distribuição de Gauss cuja largura é, essencialmente, função da estatística de multiplicação do 1.º dínodo do PM2. Na fig. 3 apresenta-se um espectro-tipo para uma produção de um número médio de fotoelectrões $n_{fe} \simeq 4$.

Como referimos anteriormente, foram estudadas fibras de diferentes características. Na fig. IV.4 estão representados os resultados obtidos com fibras azuis e verdes de diâmetro $\phi = 1$ mm e 0.5 mm.

Estes resultados permitem-nos deduzir que:

i) a produção de luz nas fibras de 0.5 mm comparada à das fibras de 1 mm é razoável sem que haja uma degradação no comprimento de atenuação;

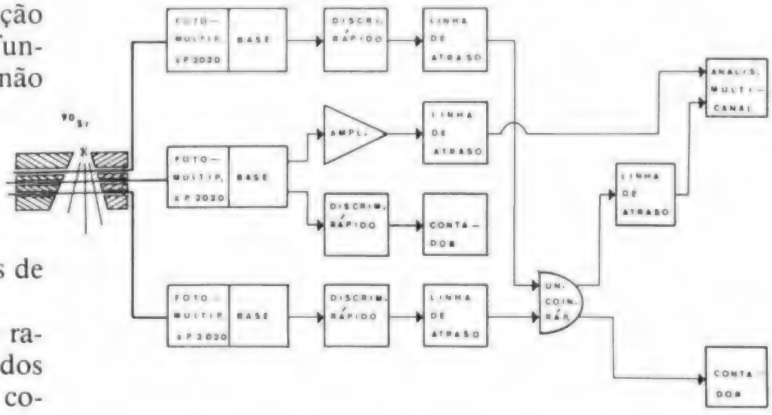


Fig. IV.2 Dispositivo experimental para medição do número médio de fotoelectrões n_{fe} e do comprimento de atenuação da fibra óptica cintilante F2.

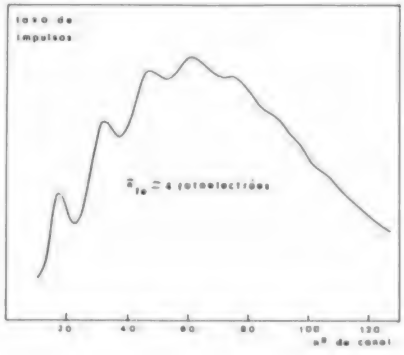


Fig. IV.3 Espectro de fotoelectrões observados no selector de amplitude multicanal correspondente a um número médio de fotoelectrões $n_{fe} \simeq 4$.

ii) as fibras verdes, apresentam um comprimento de atenuação análogo ao das fibras azuis ($L_{at} \sim 1.50$ m).

No que respeita à resistência à radiação nuclear, resultados preliminares permitiram observar que, quando sujeitas a elevadas doses de radiação γ (1 a 30 Mrad em 3 a 10 dias) a degradação é mais pronunciada para as fibras azuis (Popop) do que para as fibras verdes (K27).

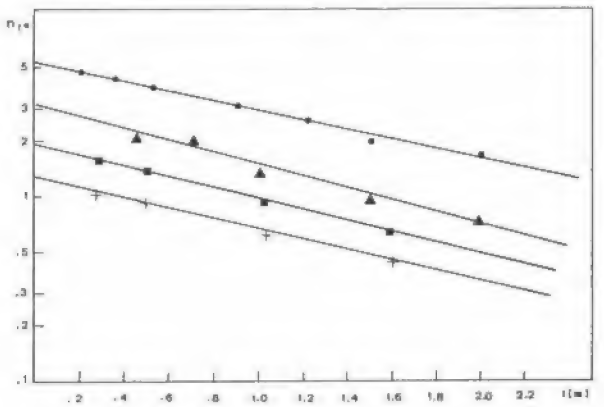


Fig. IV.4 Número médio de fotoelectrões por electrão, n_{fe}, em função da distância *l* a que a luz é transmitida.
fibra azul de diâmetro ϕ = 1 mm
fibra verde de diâmetro ϕ = 1 mm
2 fibras azuis de diâmetro ϕ = 0,5 mm
1 fibra azul de diâmetro ϕ = 0,5 mm

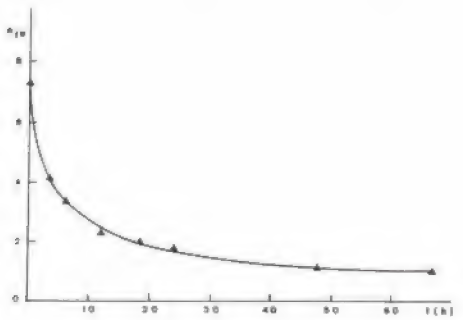


Fig. IV.5 Número de fotoelectrões por electrão atravessando uma fibra de 1 mm de diâmetro em função do tempo (t) de exposição a uma fonte intensa de U.V.

No desenvolvimento da técnica de utilização da própria fibra cintilante como seu guia de luz, foram irradiadas diversas fibras com radiação U.V. e foi medido o número médio de fotoelectrões produzidos em função do número de horas de irradiação (t). A dose de irradiação recebida pela fibra é proporcional a t. Na fig. 5 apresentam-se os resultados obtidos com 6 lâmpadas Philips TL de 40 Watt. No que respeita à degradação da produção de luz em função do tempo de exposição, resultados análogos foram obtidos com fontes mais intensas (taxa de dose mais elevada) e menos intensas.

Os resultados preliminares obtidos, permitem-nos prosseguir com o estudo da degradação das propriedades ópticas e de cintilação das fibras ópticas cintilantes de plástico, quando sujeitas à radiação nuclear ou à radiação U.V. embora por razões diversas como indicámos anteriormente.

BIBLIOGRAFIA

NA10:
L. Anderson et al., 223 (1984) 26
B. Betev et al., Z. Phys. C 28 (1985) 15
B. Betev et al., Z. Phys. C 28 (1985) 9

NA38:
A. Baldit et al., Proposal CERN/SPSC/P211
A. Baldit et al., Addendum to proposal CERN/SPSC/P211

WA74:
T. Armstrong et al. Proposal CERN/SPCS/81-/0.6
T. Armstrong et al. NPB262 (85)356

FIBRAS ÓPTICAS CINTILANTES DE PLÁSTICO:
Optical scintillating fibres for particle detectors
L.R. Allemanad et al.
NIM 225 (1984) 522

Electromagnetic calorimetry using scintillating plastic fibres
H. Burmeister et al.
NIM 225 (1984) 530.

TEORIA E PROPRIEDADES DE PLASMAS PRODUZIDOS POR ONDAS DE SUPERFÍCIE

Carlos Matos Ferreira
Departamento de Física e Centro de Electrodinâmica (I.N.I.C.)

RESUMO

Analisam-se neste trabalho os processos básicos que ocorrem numa coluna de plasma criada e mantida por uma onda de superfície. Desenvolve-se um modelo teórico que permite determinar de forma consistente as distribuições radiais da densidade do plasma e do campo eléctrico da onda, bem como a distribuição de densidade ao longo da coluna, em função das condições de funcionamento. Mostra-se que existem leis de semelhança para este tipo de descarga que permitem a extensão dos resultados obtidos a condições experimentais diversificadas. O bom acordo obtido entre teoria e experiência permite concluir que o estudo destes plasmas abre o caminho para uma compreensão mais profunda das descargas HF.

ABSTRACT

This paper discusses the basic processes occurring in a plasma column sustained by a surface wave. The theoretical model here developed treats in a self-consistent way the radial distributions of the plasma density and of the SW field intensity and provides detailed predictions concerning the radial and the axial variations in the plasma properties as a function of the discharge parameters. It also predicts the existence of simple similarity laws which enable the extension of the results to different experimental circumstances. The theory is shown to agree well with the experiments. Hence, the study of these discharges opens the way to a deeper understanding of HF plasmas in general.

1. INTRODUÇÃO

Designam-se genericamente por ondas de superfície (O. S.) as ondas que se propagam ao longo da superfície de separação entre dois meios sem que haja conversão de energia do campo associado com a onda para qualquer outro modo (sem radiação). A designação é apropriada já que, por um lado, a energia da onda é essencialmente transportada numa zona estreita junto da interface e que, por outro, a onda segue ao longo da interface, mesmo quando esta é encurvada, como se estivesse ligada a ela.

Lord Rayleigh foi o primeiro a estabelecer que se podem excitar, através duma fonte de dimensões finitas, dois tipos de ondas num meio elástico semi-infinito: ondas espaciais (também designadas por ondas radiadas) que se propagam em todas as direcções; ondas de superfície que se propagam ao longo da fronteira.

A situação é análoga no caso das ondas electromagnéticas. Se considerarmos, por exemplo, um cilindro fino, infinitamente longo, de material dieléctico sem perdas, imerso no ar, é possível a propagação duma onda electromagnética de superfície ao longo da interface. A intensidade quer do campo eléctrico, quer do campo magnético, da onda decai radialmente, para uma distância r suficientemente grande, segundo a forma $r^{-1/2}$ exp $(-k_\perp r)$, sendo $k_\perp$ a constante de decaimento. Na ausência de perdas $k_\perp$ é real, pelo que não há variação de fase ao longo de r; não existe, portanto, fluxo de energia perpendicularmente à interface, mas apenas ao longo dela.

O decaímento do tipo exponencial da intensidade dos campos a partir da interface, para um ou ambos os lados, é uma característica das ondas de superfície. No caso dos dieléctricos usuais (isto é, excluindo os plasmas) só existe decaímento exponencial no dieléctrico de menor permitividade; no outro, a intensidade dos campos oscila espacialmente (é, nomeadamente, o que acontece no interior do cilindro de dieléctrico no exemplo considerado acima).

Contudo, se um dos meios for, por exemplo, um plasma frio e sem perdas, a sua permitividade

$$\epsilon_p = 1 - \omega_{pe}^2 / \omega^2$$

Original recebido para publicação em 8/6/85.

poderá ser negativa, o que se verifica quando a frequência ω for inferior à frequência plasma electrónica ω_{pe}. Trata-se do único dieléctrico conhecido com esta propriedade. Na superfície de separação de tal meio com um dieléctrico convencional de permitividade $\epsilon_g > 0$ pode propagar-se uma O. S. com campos evanescentes em ambos os lados da interface, desde que se verifique $|\epsilon_p| > \epsilon_g$[1]. O carácter evanescente puro apenas se observará, no entanto, se ambos os meios forem semi-infinitos. No caso dum plasma de dimensões transversais finitas (por exemplo, coluna cilíndrica, fatia de plasma de geometria plana, etc.) o campo no plasma não é propriamente evanescente, embora a sua intensidade decresça da fronteira para o eixo. As O. S. com estas características chamam-se ondas de superfície puras*[2, 3] e serão as únicas consideradas neste artigo.

O estudo das O. S. em plasmas, desde a sua descoberta no trabalho pioneiro de Trivelpiece[4] em 1958, tem motivado numerosos trabalhos teóricos e experimentais (veja-se o excelente artigo de revisão publicado recentemente por Moisan et al[5]). As razões dum tal interesse foram, inicialmente, problemas de natureza fundamental em que intervêm as O. S., tais como o aquecimento de plasmas por campos electromagnéticos, a absorção de radiação laser por plasmas densos, as interacções não lineares como, por exemplo, as instabilidades paramétricas. Outro motivo de interesse residia no facto das O. S. poderem ser usadas para diagnóstico de plasmas de baixa temperatura (por exemplo, plasmas de descarga) permitindo medir a densidade, a frequência de colisão e a velocidade de deriva dos electrões[6-9]. Em todos estes casos, no entanto, o plasma é suposto ser criado por mecanismos independentes da O. S., ou seja, esta propaga-se num plasma que se supõe pré-existente.

Em anos mais recentes, no entanto, tem vindo a desenvolver-se um novo campo de investigação baseado na perspectiva inovadora da criação e manutenção de plasmas por ondas de superfície[10-13]. Este tipo de plasmas, que designaremos abreviadamente por plasmas de O. S., podem ser facilmente produzidos em laboratório, desde pressões da ordem dos militorr até algumas atmosferas, usando quaisquer frequências desde alguns MHz até cerca de 10 GHz, o que permite cobrir todo o espectro desde as radiofrequências (rf) às micro-ondas.

A densidade do plasma de O. S. pode variar tipicamente entre 10^{10}cm^{-3} e 10^{14}cm^{-3}, consoante a potência HF absorvida, a pressão do gás e o diâmetro do tubo de descarga. Por exemplo, pode-se produzir uma coluna de plasma com cerca de 2 m de comprimento, 2,5 cm de diâmetro e uma densidade média de cerca de 10^{11}cm^{-3}, a uma pressão de alguns dezenas de militorr, usando 80 W de potência HF à frequência de 500 MHz. Os plasmas de O. S. são estáveis, calmos e perfeitamente reproductíveis, o que os torna substitutos interessantes dos plasmas de descarga convencionais (coluna positiva da descarga dc, descarga rf do tipo capacitivo, plasma produzido numa cavidade micro-onda, etc.) para aplicações de grande actualidade como o tratamento de superfícies, a gravura por plasma, a síntese química, a excitação de lasers, lâmpadas espectrais, etc.

Para além do interesse destas aplicações, potenciais ou demonstradas, acresce que o estudo dos plasmas de O. S. tem contribuído significativamente para uma compreensão mais aprofundada dos mecanismos básicos de funcionamento das descargas HF em geral (veja-se, por exemplo, sobre este assunto o artigo de Ferreira e Moisan[14]). Tal facto deve-se, por um lado, à extrema versatilidade já mencionada de se poder varrer todo o espectro, desde as radiofrequências às micro-ondas, para se criar o plasma e, por outro, à possibilidade de relacionar localmente as propriedades do plasma com a potência absorvida (possibilidade que resulta do facto do plasma ser criado, neste caso, por uma onda progressiva). Graças a estas características, a teoria dos plasmas de O. S. atingiu já um grau de maturidade satisfatório, permitindo interpretar e explicar com boa precisão os resultados experimentais.

Constitui objectivo central deste artigo discutir as propriedades dos plasmas de O. S. e os processos básicos que asseguram a sua manutenção. Por não terem ocorrido, entretanto, progressos relevantes, parte do material científico aqui apresentado é, na sua essência, embora com variações de forma, o que se tomou como base para um curso dado sobre este assunto pelo autor no NATO Advanced Study Institute on Radiative Processes in Discharge Plasmas que teve lugar no Verão de 1985[15].

2. PROPRIEDADES DAS ONDAS DE SUPERFÍCIE NUM PLASMA

2.1 Equações dos Campos

Considere-se uma O. S. propagando-se ao longo da interface (eixo $-z$) dum plasma com um dieléctrico de permitividade ϵ_g e admitimos que os campos $\vec{E}$ e $\vec{H}$ variam segundo z e no tempo segundo a forma exp j ($\omega t - \kappa z$), em que $\kappa = \beta - j\alpha$, sendo $\beta = 2\pi/\lambda$ o número de onda e α o coeficiente de atenuação.

Descrevendo o plasma como um dialéctrico de permitividade ϵ_p, as amplitudes complexas dos campos obedecem às equações

$$\vec{\nabla} \times \vec{E} = -j\,\omega\mu_o\,\vec{H} \tag{1}$$

$$\vec{\nabla} \times \vec{H} = j\,\omega\epsilon_o\epsilon_p\,\vec{E} \tag{2}$$

$$\vec{\nabla} \cdot \vec{H} = 0 \tag{3}$$

$$\vec{\nabla} \cdot (\epsilon_p\,\vec{E}) = 0 \tag{4}$$

* Quando um dos meios em presença é um plasma com $\epsilon_p < o$ e existe um campo magnético aplicado paralelo à fronteira podem propagar-se igualmente ondas de superfície cuja intensidade do campo oscila espacialmente no interior do plasma.

A permitividade do plasma é, em geral, um número complexo. Está relacionada com a conductividade complexa do plasma, σ, através da expressão

$$\epsilon_p = 1 - j\frac{\sigma}{\omega\epsilon_o}, \tag{5}$$

sendo σ dado por[16]

$$\sigma = -\frac{2}{3}\frac{e^2 n_e}{m}\int_0^\infty \frac{u^{3/2}}{\upsilon_c + j\omega}\frac{df}{du}\,du \tag{6}$$

Nestas expressões e e m representam a carga e a massa do electrão, ϵ_o é a constante dieléctrica do vácuo, n_e é a densidade electrónica, υ_c é a frequência de colisão electrão-neutrão para a transferência de quantidade de movimento, $u = mv^2/2e$ é a energia do electrão expressa em volts e $f(u)$ é a função de distribuição de energia (FDE) dos electrões com a normalização $\int_0^\infty f(u)\,u^{1/2}\,du = 1$. Quando υ_c é independente de u a Eq. (6) reduz-se à forma $\sigma = e^2 n_e/m(\upsilon_c + j\omega)$; por outro lado, na ausência de colisões ($\upsilon_c = 0$) σ é imaginário puro e a permitividade do plasma reduz-se à forma

$$\epsilon_p = 1 - (\omega_{pe}/\omega)^2, \tag{7}$$

em que $\omega_{pe} = (e^2 n_e/\epsilon_o m)^{1/2}$ é a frequência angular plasma electrónica.

Tendo em atenção que $\vec{\nabla} \equiv \vec{\nabla}_T - j\kappa\,\vec{e}_z$, em que $\vec{e}_z$ é o vector unitário do eixo $\overline{zz}$ e $\vec{\nabla}_T$ é a parte do operador $\vec{\nabla}$ que envolve unicamente as derivadas em relação às coordenadas transversais, obtém-se facilmente das eqs. (1) e (2)

$$\vec{\nabla}_T \times \vec{E}_T = -j\omega\mu_o\,\vec{H}_z \tag{8}$$

$$\vec{\nabla}_T \times \vec{E}_z - j\kappa\,\vec{e}_z \times \vec{E}_T = -j\omega\mu_o\,\vec{H}_T \tag{9}$$

$$\vec{\nabla}_T \times \vec{H}_T = j\omega\,\epsilon_o\,\epsilon_p\,\vec{E}_z \tag{10}$$

$$\vec{\nabla}_T \times \vec{H}_z - j\kappa\,\vec{e}_z \times \vec{H}_T = j\omega\,\epsilon_o\,\epsilon_p\,\vec{E}_T, \tag{11}$$

em que $\vec{E}_T$ e $\vec{H}_T$ designam as componentes transversais de $\vec{E}$ e $\vec{H}$.

Multiplicando vectorialmente a eq. (9) por $\vec{e}_z$ e combinando com (11) obtém-se facilmente

$$\vec{E}_T = -\frac{j\kappa}{\frac{\omega^2}{c^2}\epsilon_p - \kappa^2}\vec{\nabla}_T E_z + \frac{j\omega\mu_o}{\frac{\omega^2}{c^2}\epsilon_p - \kappa^2}(\vec{e}_z \times \vec{\nabla}_T H_z), \tag{12}$$

em que $c = (\epsilon_o\,\mu_o)^{-1/2}$ é a velocidade da luz no vácuo.

Procedendo de igual forma com as eqs. (8) e (10) obtém-se agora

$$\vec{H}_T = -\frac{j\kappa}{\frac{\omega^2}{c^2}\epsilon_p - \kappa^2}\vec{\nabla}_T H_z - \frac{j\omega\,\epsilon_o\,\epsilon_p}{\frac{\omega^2}{c^2}\epsilon_p - \kappa^2}(\vec{e}_z \times \vec{\nabla}_T E_z). \tag{13}$$

As eqs. (12) e (13) constituem expressões gerais, relacionando as componentes transversais $\vec{E}_T$ e $\vec{H}_T$ com os gradientes transversos das componentes longitudinais, para qualquer onda progressiva. Reduzem-se a formas mais simples no caso de ondas TE ($E_z = 0$), ou de ondas TM ($H_z = 0$). Neste último caso, que nos interessará particularmente no seguimento, obtém-se simplesmente

$$\vec{E}_T = -\frac{j\kappa}{\frac{\omega^2}{c^2}\epsilon_p - \kappa^2}\vec{\nabla}_T E_z \tag{14}$$

$$\vec{H}_T = \frac{\omega\,\epsilon_o\,\epsilon_p}{\kappa}(\vec{e}_z \times \vec{E}_T). \tag{15}$$

No caso dum plasma homogéneo (ϵ_p = const.) as equações que permitem obter E_z e H_z são particularmente simples.

Tendo em conta a identidade $\vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{A}) - \nabla^2\vec{A}$, obtém-se facilmente a partir das eqs. (1)-(3)

$$\nabla_T^2\,\vec{E} + \left(\frac{\omega^2}{c^2}\epsilon_p - \kappa^2\right)\vec{E} = 0 \tag{16}$$

$$\nabla_T^2\,\vec{H} + \left(\frac{\omega^2}{c^2}\epsilon_p - \kappa^2\right)\vec{H} = 0 \tag{17}$$

portanto, E_z e H_z satisfazem às equações

$$\nabla_T^2\,E_z + \left(\frac{\omega^2}{c^2}\epsilon_p - \kappa^2\right)E_z = 0 \tag{18}$$

$$\nabla_T^2\,H_z + \left(\frac{\omega^2}{c^2}\epsilon_p - \kappa^2\right)H_z = 0. \tag{19}$$

A hipótese de homogeneidade do plasma é, no entanto, inconsiste na generalidade dos casos e conduz a erros significativos. Por exemplo, se considerarmos uma coluna de plasma criada e mantida por uma O. S. em condições de atenuação fraca ($\omega \gg \upsilon_c$), é razoável desprezar a inomogeneidade segundo o eixo do tubo, mas há que ter em conta, de forma consistente, a inomogeneidade nas direcções transversais. Neste caso, partindo da Eq. (4) obtém-se $\vec{\nabla} \cdot \vec{E} = -\vec{E} \cdot \vec{\nabla}\epsilon_p/\epsilon_p$, pelo que a Eq. (16) adquire a forma mais geral

$$\nabla_T^2\,\vec{E} + \left(\frac{\omega^2}{c^2}\epsilon_p - \kappa^2\right)\vec{E} + \vec{\nabla}\left(\vec{E}_T \cdot \frac{\vec{\nabla}_T\,\epsilon_p}{\epsilon_p}\right) = 0. \tag{20}$$

Para uma onda TM $\vec{E}_T$ é dado pela Eq. (14), a qual combinada com a Eq. (20) permite obter, neste caso, a seguinte expressão para E_z

$$\nabla_T^2 E_z + \left(\frac{\omega^2}{c^2}\epsilon_p - \kappa^2\right) E_z - \frac{\kappa^2}{\frac{\omega^2}{c^2}\epsilon_p - \kappa^2} \vec{\nabla}_T E_z \cdot \frac{\vec{\nabla}_T \epsilon_p}{\epsilon_p} = 0 \tag{21}$$

De forma análoga, na Eq. (17) aparece um termo adicional escrevendo-se, agora, esta na forma

$$\nabla_T^2 \vec{H} + \left(\frac{\omega^2}{c^2}\epsilon_p - \kappa^2\right) \vec{H} + j\,\omega\,\epsilon_o\,\vec{\nabla}_T\,\epsilon_p \times \vec{E} = 0 \tag{22}$$

Esta equação, conjuntamente com a Eq. (12), determina H_z.

2.2 Ondas TM propagando-se ao longo dum plasma cilíndrico

As ondas de superfície podem expressar-se, em geral, como sobreposição de modos TE e TM. Contudo, é conhecido que apenas os modos TM podem satisfazer às condições de fronteira apropriadas, na interface plasma-dieléctrico, no caso de ondas com simetria axial. Assim, por exemplo, em geometria cilíndrica o único modo axialmente simétrico é o modo TM com $m = 0$, em que m é o número de onda correspondente ao ângulo azimutal ϕ; neste caso só existem as componentes E_r, E_z e H_ϕ. Todos os modos com $m \geqslant 1$ são híbridos, possuindo simultaneamente componentes E_z e H_z não nulas.

As equações dos campos do modo TM no plasma obtêm-se facilmente aplicando as equações gerais (14), (15) e (21):

$$\frac{d^2 E_z}{dr^2} + \left(\frac{1}{r} - \frac{\kappa^2}{\frac{\omega^2}{c^2}\epsilon_p - \kappa^2}\frac{1}{\epsilon_p}\frac{d\epsilon_p}{dr}\right)\frac{dE_z}{dr} + \left(\frac{\omega^2}{c^2}\epsilon_p - \kappa^2\right) E_z = 0 \tag{23}$$

$$E_r = -j\,\frac{\kappa}{\frac{\omega^2}{c^2}\epsilon_p - \kappa^2}\,\frac{dE_z}{dr} \tag{24}$$

$$H_\phi = \frac{\omega\,\epsilon_o\,\epsilon_p}{\kappa}\,E_r \tag{25}$$

Este modo é facilmente excitado usando um dispositivo HF especial designado por "Surfatron"[10, 12]. Tem sido este o modo usado, em geral, para a criação de plasmas de O. S., pelo que nos limitaremos a este caso no seguimento.

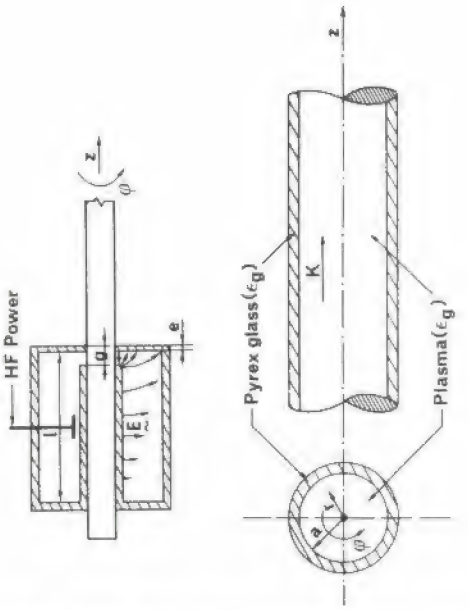


Fig. 1 Diagrama esquemático do dispositivo HF e do acoplador num Surfatron e da geometria do tubo de descarga

A Fig. 1 representa esquematicamente uma instalação experimental usada para produzir um plasma num tubo de vidro pyrex através duma O. S. "lançada" por um "Surfatron".

As características de dispersão das O. S. numa tal geometria obtêm-se a partir da continuidade das componentes E_z e H_ϕ nas interfaces plasma-vidro e vidro-ar (isto é, para $r = a$ e $r = b$, sendo a e b os raios interior e exterior do tubo, respectivamente). Para tal, basta ter em conta as Eqs. (23)-(25) do campo no plasma e equações similares quer no vidro, quer no ar exterior (com ϵ_p substituído por ϵ_g, constante, no vidro e por $\epsilon_{ar} = 1$, no ar).

O cálculo das características de dispersão só é simples, no entanto, no caso de ondas fracamente atenuadas ($\alpha \ll \beta$), o que permite usar álgebra real, e admitindo que o plasma é homogéneo[11, 17, 18]. Neste caso, as componentes E_z e E_r no plasma são dadas pelas equações simples.

$$E_z(r) = E(o)\,I_0(\Gamma r), \tag{26}$$

$$E_r(r) = E(o)\,I_1(\Gamma r), \tag{27}$$

em que I_0 e I_1 são as funções de Bessel modificadas de ordem zero e um, respectivamente, E(o) é a intensidade do campo no eixo, e Γ é definido por

$$\Gamma^2 = \beta^2 + \frac{\omega_{pe}^2 - \omega^2}{c^2} \tag{28}$$

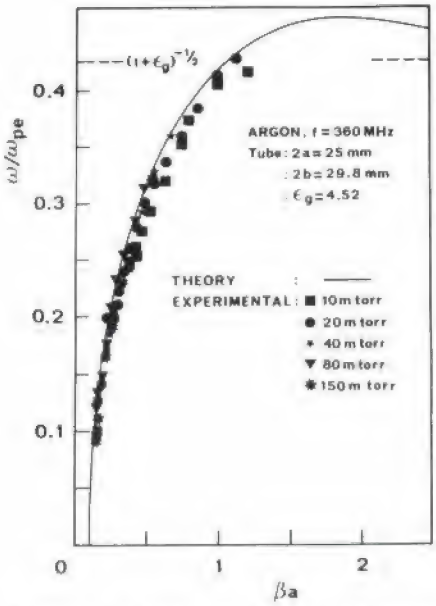


Fig. 2 Características de dispersão duma onda de superfície com f = 360 MHz num plasma de argon

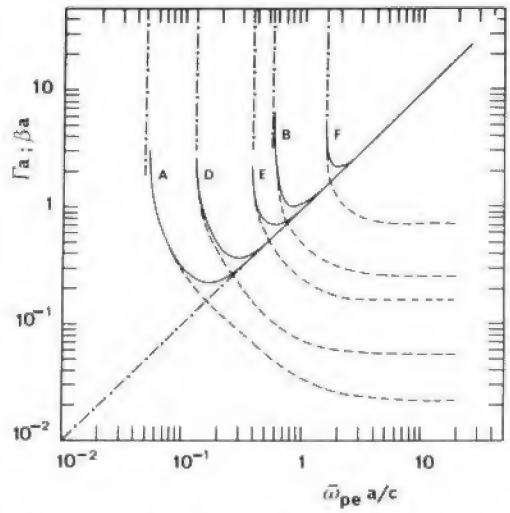


Fig. 3 Curvas de βa (tracejado) e Γa (cheio) vs. ω_{pe} a/c num plasma de argon homogéneo: a = 0,45 cm, b = 0,60 cm, f = 210 MHz (A) e 2450 MHz (B); a = 1,25 cm, b = 1,49 cm, f = 210 MHz (D), 600 MHz (E), e 2450 MHz (F)

A Fig. 2 mostra características de dispersão calculadas segundo este modelo simplificado e resultados experimentais obtidos (experimentalmente estes resultados são obtidos com ω fixo e fazendo variar ω_{pe}, pelo que o diagrama da figura se designa geralmente por diagrama de fase). Apesar das hipóteses simplificadoras usadas no cálculo, o acordo entre teoria e experiência é plenamente satisfatório, excepto para βa elevado (o que é devido à não se ter considerado a inomogeneidade radial do plasma, como veremos em breve).

Observam-se os seguintes limites no diagrama de fase. Quando $\omega/\omega_{pe} \to 0$ verifica-se que $\beta \to (\omega/c)\sqrt{\epsilon}$, em que $1<\epsilon<\epsilon_g$. Significa isto que, quando $\omega_{pe} >> \omega$, o número de onda da O. S. é próximo do que se observaria para uma onda plana electromagnética propagando-se num dieléctrico de permitividade ϵ. No limite oposto, $\beta \to \infty$, verifica-se que $\omega/\omega_{pe} \to (1+\epsilon_g)^{-1/2}$, ou seja que $|\epsilon_p| \to \epsilon_g$. Este limite corresponde ao corte: a O. S. não se pode propagar se a densidade do plasma for inferior a um certo valor crítico, n_c, dado pela relação $|\epsilon_{pc}| = \epsilon_g$. Numericamente, tem-se

$$n_c\,(\mathrm{cm}^{-3}) = 1{,}245 \times 10^{-8}\, f^2 (1+\epsilon_g), \qquad (29)$$

com f expresso em Hertz.

A Fig. 3 mostra os valores de Γa e de βa, calculados segundo este modelo, em função de ω_{pe} a/c, para vários valores da frequência e dos parâmetros geométricos a e b. Na vizinhança do corte tem-se $\Gamma \simeq \beta$ e Γa torna-se, por isso, elevado.

As Eqs. (26) e (27) revelam então que a intensidade do campo é muito maior junto da interface do que na zona axial do tubo. No limite oposto ($\omega_{pe} >> \omega$, isto é, densidade do plasma elevada) verifica-se uma situação semelhante, já que, nesse caso, se tem $\beta << \Gamma$ e $\Gamma a \to \omega_{pe}$ a/c>>1.

É importante notar que, quando a densidade do plasma é suficientemente elevada ($\omega_{pe} >> \omega$), o quociente E(a)/E(o) é tanto maior quanto maior for a densidade do plasma; isto é, o campo concentra-se junto da interface plasma-vidro à medida que a densidade aumenta. Esta conclusão é igualmente válida para um plasma radialmente inomogéneo (o que conta nesse caso é o valor da densidade média numa fatia de plasma) como veremos no seguimento.

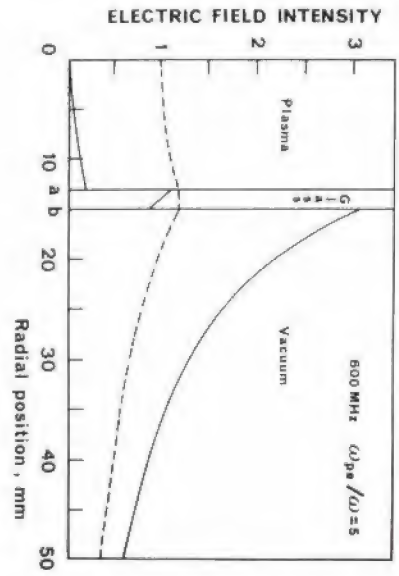


Fig. 4 Distribuição radial das componentes radial (cheio) e axial (tracejado) do campo eléctrico duma O. S. TM propagando-se ao longo duma coluna cilíndrica de plasma homogéneo [da ref.[5]]

A Fig. 4 mostra o perfil radial das componentes E_z e E_r no plasma, no vidro e no ar exterior, no caso $f = 600$ MHz e $\omega_{pe}/\omega = 5$.

O efeito da inomogeneidade radial do plasma sobre o diagrama de fase da O. S. tem sido investigado por diversos autores, usando a aproximação quase-estática ($\vec{\nabla} \times \vec{E} = 0 \rightarrow \vec{E} = -\vec{\nabla} F$). Trivelpiece[19] e Ilić[20] partiram da hipótese dum perfil parabólico da densidade, do tipo

$$n_e(r) = n_e(0)\,[1 - K(\frac{r}{a})^2], \qquad (30)$$

em que $0 \leqslant K < 1$ constitui um parâmetro de forma do perfil. Os diagramas de fase $\omega/\bar{\omega}_{pe}$ vs. βa obtidas por Ilić, para diversos valores de K, são representados na Fig. 5 [($\bar{\omega}_{pe}$ corresponde à densidade média $\bar{n}_e = n_e(0)(1 - K/2)$]. Verifica-se que para $\beta a < 1$ o efeito da inomogeneidade é desprezável, tudo se passando como se a O. S. se propagasse num plasma homogéneo com densidade $\bar{n}_e$. Contudo, há diferenças significativas entre os dois casos quando $\beta a > 1$. A razão disso é, essencialmente, o facto do campo se concentrar, nesse caso, junto da interface, pelo que o corte fica determinado pela densidade na fronteira $n_e(a)$ e não pela densidade média (ou seja, $\omega/\omega_{pe}(a) \rightarrow (1+\epsilon_g)^{-1/2}$ quando $\beta a \rightarrow \infty$). Sendo assim, obtém-se $\omega/\bar{\omega}_{pe} \rightarrow (1+\epsilon_g)^{-1/2}\,[(1-K)/(1-K/2)]^{1/2}$ quando $\beta a \rightarrow \infty$, valor que é sempre inferior ao que corresponde a um plasma homogéneo com o mesmo valor de $\bar{\omega}_{pe}$.

Nos diagramas da Fig. 5 aparecem zonas de propagação retrógrada que podem, no entanto, como mostrou Ilić[20], ser eliminadas considerando ainda o efeito duma baínha junto da parede do tubo. Ilić apenas considerou, contudo, o caso duma baínha sem electrões e dum plasma de densidade acima do valor crítico na fronteira plasma-baínha, não tendo, portanto, analisado os efeitos da ressonância que ocorre quando $\omega_{pe} = \omega$. Ora, se a densidade dos electrões decresce regularmente na baínha até se anular na parede, uma tal ressonância deverá ocorrer algures, o que produzirá amortecimento dos modos de O. S. por excitação de ondas de Langmuir na região de ressonância[21]. Este amortecimento virá adicionar-se ao produzido pela colisões electrão-neutro, colisões devido às quais os electrões do plasma extraem energia da onda. No momento actual, a teoria dos plasmas de O. S. apenas tem em conta o amortecimento colisional, o que se tem revelado suficiente para a interpretação da generalidade dos resultados experimentais.

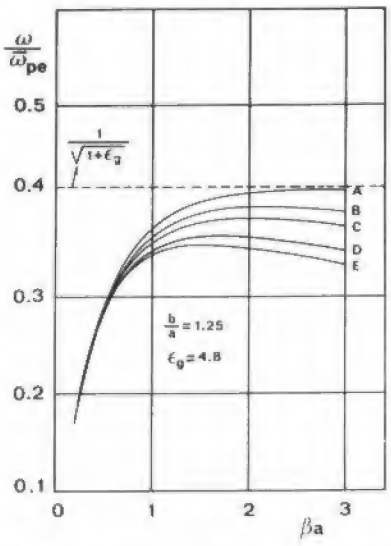


Fig. 5 Efeito dum perfil de densidade parabólico $n_e = n(o)\,[1 - K(r/a)^2]$ sobre as características de dispersão do modo TM, para k = O (A); 0,4 (B); 0,6 (C); 0,8 (D); e 0,9 (E) [da ref.[20]]

A potência extraída da onda por uma unidade de comprimento, ao longo da direcção de propagação, devido às colisões elásticas electrão-neutro, é dada por

$$\frac{dP_{abs}}{dz} = \frac{1}{2}\int_0^a R_e(\sigma)\,E^2\,2\pi r\,dr, \qquad (31)$$

em que $R_e(\sigma)$ é a parte real da conductividade, dada pela Eq. (6), e $E^2 = E_z^2 + E_r^2$. Em regime fracamente colisional (atenuação fraca: $\nu_c \ll \omega$) a Eq. (6) pode-se escrever aproximadamente na forma

$$\mathrm{Re}(\sigma) = \frac{e^2 n_e}{m\omega^2}\,\bar{\nu}_c, \qquad (32)$$

em que se definiu a frequência de colisão efectiva

$$\bar{\nu}_c = -\frac{2}{3}\int_0^\infty u^{3/2}\,\nu_c\frac{df}{du}\,du. \qquad (33)$$

Obtém-se assim

$$\frac{dP_{abs}}{dz} = \frac{e^2}{2m}\int_0^a \bar{\nu}_c(\frac{E}{\omega})^2\,2\pi r\,dr, \qquad (34)$$

expressão que, no caso dum plasma homogéneo, se poderá escrever ainda na forma

$$\frac{dP_{abs}}{dz} = \frac{\bar{\nu}_c\epsilon_o}{2}(\frac{\omega_{pe}}{\omega})^2\int_0^a E^2\,2\pi r\,dr. \qquad (35)$$

O coeficiente de atenuação (colisional) da onda pode então calcular-se a partir da expressão (22)

$$\alpha = \frac{1}{2}\,\frac{dP_{abs}/dz}{v_g w_t}, \qquad (36)$$

em que v_g designa a velocidade de grupo e w_t a média no tempo da energia total (eléctrica + magnética + cinética das partículas carregadas do plasma) por unidade

de comprimento, na direcção da propagação. Num plasma frio e em condições de fraca atenuação, $v_g\ w_t$ é igual ao fluxo total de energia electromagnética através dum plano perpendicular à direcção de propagação, ou seja, igual ao fluxo total da parte real do vector de Poynting complexo $\vec{E} \times \vec{H}^*/2$.

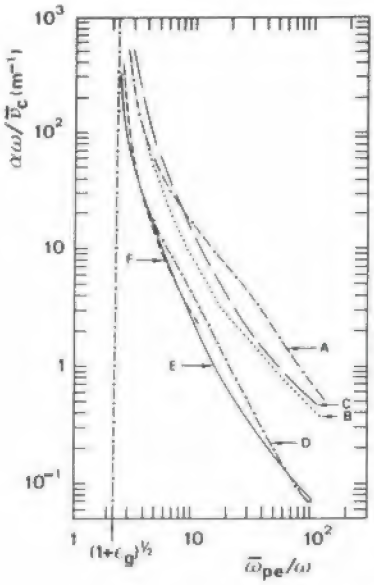


Fig. 6 Características de atenuação duma O. S. TM numa coluna cilíndrica de plasma de argon homogéneo: a = 0,45 cm, b = 0,60 cm, f = 210 MHz (A) e 2450 MHz (B); a = 0,30 cm, b = 0,45 cm, f = 2450 MHz (C); a = 1,25 cm, b = 1,49 cm, f = 210 MHz (D), 600 MHz (E) e 2450 MHz (F)

A Fig. 6 mostra valores de $\alpha\omega/\bar{v}_c$ em função de ω_{pe}/ω, calculados para um plasma homogéneo a partir do teorema de Poynting. Estes cálculos simplificados, que envolvem unicamente álgebra real, estão em excelente acordo com os resultados mais precisos que se podem obter a partir das soluções da equação de dispersão em álgebra complexa[23]. É de notar, a partir dos resultados da Fig. 6, que a atenuação é pequena quando $\omega_{pe} >> \omega$ (densidades do plasma elevadas), mas que ela se torna extremamente elevada na vizinhança da densidade de corte (convém ter em atenção que, na situação experimental que temos vindo a analisar, a densidade de corte corresponde aproximadamente ao fim da coluna de plasma; a densidade do plasma decresce regularmente ao longo da direcção de propagação, desde a saída do surfatron até se atingir a condição de corte, o que determina o fim da coluna).

3. PROPRIEDADES DO PLASMA MANTIDO POR UMA ONDA DE SUPERFÍCIE

3.1 Processos básicos

Os processos básicos de manutenção dum plasma por um campo eléctrico HF, em regime de absorção colisional, são simples de entender.

Devido às colisões electrão-neutro, o movimento dos electrões tem uma componente em fase com o campo, pelo que os electrões absorvem energia deste. Esta energia é dissipada principalmente em colisões elásticas e inelásticas com os átomos do gás, incluindo-se nas colisões inelásticas a ionização dos átomos por impacto electrónico com produção de novos electrões livres. Para que o plasma possa, então, ser mantido pelo campo HF é necessário que a taxa de produção de novos electrões por ionização compense exactamente a taxa de perda destas partículas por difusão para as paredes e, eventualmente, por recombinação. É necessário ainda, simultaneamente, que a potência absorvida por electrão, em média no volume do plasma, seja igual à potência média perdida por electrão devido aos processos colisionais e de transporte para a parede circundante.

A primeira destas condições pode ser adequadamente formulada através dum modelo do tipo hidrodinâmico, envolvendo as equações de continuidade e de transporte de quantidade de movimento, na presença do campo ambipolar, para os electrões e os iões. A segunda pode ser formulada em termos cinéticos, isto é, através da própria equação de Boltzmann, se admitirmos que os electrões se encontram localmente em equilíbrio com o campo eléctrico HF e que os gradientes de difusão são suficientemente pequenos para não afectarem a distribuição de energia dos electrões. Começaremos por analisar a equação de Boltzmann e o balanço local de energia.

3.2 Função de distribuição de energia (FDE) dos electrões. Balanço local de energia

A situação que analisaremos aqui é a dum gás de electrões na presença dum campo HF, $E = E_p \exp(j\omega t)$[24]. Admitiremos, por hipótese, que: *a*) A amplitude do campo e os gradientes de difusão são suficientemente pequenos para que a função de distribuição de velocidades dos electrões possa ser representada por um desenvolvimento em harmónicas esféricas limitado aos dois primeiros termos (parte isotrópica + uma anisotropia de 1.ª ordem). Tal hipótese requer, em particular, que a secção eficaz total de difusão seja muito maior que a de colisão inelástica e que o livre percurso médio dos electrões seja muito inferior a qualquer das dimensões do recipiente; *b*) A frequência ω é suficientemente elevada para que a parte isotrópica da distribuição não se modifique de forma apreciável durante um ciclo de oscilação do campo. Esta hipótese é válida desde que $\omega >> \delta_e^{-1}$, em que δ_e é o tempo característico de relaxação da energia por colisões com os átomos (note-se que, para colisões elásticas, $\delta_e^{-1} \simeq (m/M)\,\upsilon_c$, sendo M a massa dos átomos, enquanto que, para as colisões ine-

lásticas, $\delta_e^{-1} \simeq \nu_x$, sendo ν_x a frequência de colisão inelástica total); *c*) Os fluxos no espaço de configuração e no de velocidades resultantes dos gradientes de difusão e do campo de carga de espaço são pequenos e podem ser desprezados, isto é, a distribuição de energia dos electrões é predominantemente determinada pelo campo HF aplicado e pelas colisões elásticas e inelásticas (uma discussão dos efeitos do campo de carga de espaço sobre a distribuição pode encontrar-se em[25, 26]).

Com estas hipóteses e seguindo o desenvolvimento feito por Holstein[27], a parte isotrópica da distribuição é dada pela equação[24] (equação de Boltzmann homogénea)

$$\frac{2}{3}\frac{d}{du}\left[u^{3/2}\,\nu_c\left(u_c\frac{df}{du}+\frac{3m}{M}f\right)\right]=\nu_x(u)\,f(u)\sqrt{u}-\sum_j \nu_j(u+V_j)\,f(u+V_j)\sqrt{(u+V_j)}, \tag{37}$$

em que

$$u_c(u)=\frac{e}{m}\,\frac{1}{\nu_c^2+\omega^2}\,\frac{E_p^2}{2} \tag{38}$$

representa a energia média extraída pelos electrões de energia u do campo aplicado, por colisão (portanto, o produto $\nu_c\,u_c$ é a potência transferida). Na Eq. (37) ν_j representa a frequência de excitação do nível j, com limiar V_j, a partir do fundamental e $\nu_x=\sum_j \nu_j$; a soma $\sum_j$ inclui a ionização, desde que se despreze a energia do electrão secundário que nasceu; a normalização de f é, como anteriormente, $\int_0^\infty f\sqrt{u}\,du=1$.

Multiplicando ambos os membros da Eq. (37) pela energia do electrão e integrando sobre todas as energias, $0\leqslant u\leqslant\infty$, obtém-se a equação de conservação da energia

$$(e\,n_e)^{-1}\,R_e(\sigma)\,\frac{E_p^2}{2}=\frac{2m}{M}\langle u\,\nu_c\rangle+\sum_j V_j\langle\nu_j\rangle, \tag{39}$$

em que σ é a conductividade, dada pela Eq. (6), e $\langle\rangle$ representa a média duma dada quantidade em toda a gama de energia. A Eq. (39) exprime simplesmente o facto que a energia absorvida do campo é igual à que é perdida por colisões elásticas (1.º termo do 2.º membro) e inelásticas (2.º termo), por unidade de tempo e em média por electrão.

A Eq. (37) mostra que a FDE só depende das variáveis exteriores E_p, ω e N (densidade do gás) através da quantidade u_c. Quando $\nu_c(u)\gg\omega$, para qualquer energia u, ter-se-á $u_c\simeq$ e $E^2/m\nu_c^2\;\alpha\,(E/N)^2$, em que $E=E_p/\sqrt{2}$ é o valor eficaz do campo; a FDE será precisamente a mesma que a que corresponderia a um campo d.c. de amplitude E. O único parâmetro independente é, neste caso, E/N. No limite oposto, $\omega\gg\nu_c(u)$, ter-se-á $u_c\simeq$ e $E^2/m\omega^2\,\alpha\,(E/\omega)^2$ e a FDE só depende do parâmetro E/ω. Em qualquer situação intermédia a FDE depende simultaneamente de E/N e de E/ω, ou, se preferirmos, de E/N (ou E/ω) e de N/ω. Como $\nu_c(u)\,\alpha\,N$, é preferível usar em vez de N/ω, que tem dimensões, um parâmetro adimensional do tipo ν_{ce}/ω, sendo ν_{ce} a frequência de colisão para os electrões duma dada energia (a qual poderá ser escolhida arbitrariamente).

Quando ν_c é independente da energia do electrão é possível reduzir a dependência da FDE a um único parâmetro, para qualquer valor de ν_c/ω, introduzindo um campo efectivo[28, 29]

$$E_e=E\,\frac{\nu_c}{(\nu_c^2+\omega^2)^{1/2}}. \tag{40}$$

Neste caso, $u_c=(e/m)\,(E_e/\nu_c)^2\,\alpha\,(E_e/N)^2$, pelo que a FDE só depende, de facto, de E_e/N para qualquer valor de ν_c/ω.

Por outro lado, o balanço da potência, expresso pela Eq. (39), pode escrever-se neste caso (em $eV\,s^{-1}$).

$$\nu_c\,u_c=\frac{e\,E_e^2}{m\,\nu_c}=\frac{2m}{M}\langle u\,\nu_c\rangle+\sum_j V_j\langle\nu_j\rangle. \tag{41}$$

O conceito de campo efectivo é poderoso (todas as grandezas médias referentes aos electrões só dependeriam de E_e/N) mas apenas interessa para modelos físicos simplificados, já que, em geral e com excepção do hélio e do hidrogénio, ν_c depende fortemente da energia do electrão.

No caso mais geral em que ν_c depende de u é interessante classificar os diferentes regimes de funcionamento através dum diagrama de E/N vs. E/ω, como se mostra na Fig. 7 para o caso do argon[24]. A linha recta

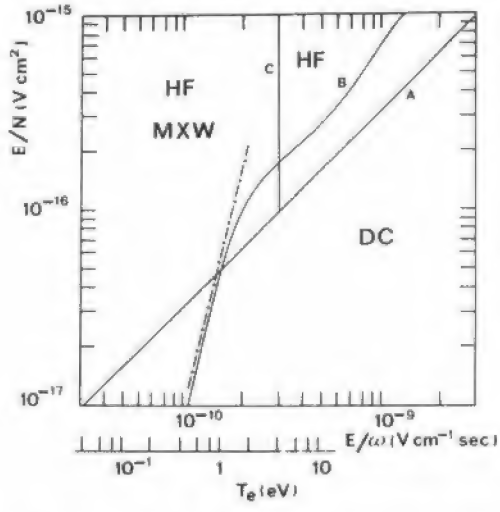


Fig. 7 Diagrama de E/N vs. E/ω: (A) $R=\omega/\nu_{cM}=1$; (B) $R'=\omega/\bar{\nu}_c=10$; (C) $E/\omega=3\times10^{-10}$ V cm⁻¹s. A curva a traço-ponto é o limite assimptótico da linha R' = 10 para pequenos valores de E/ω

A corresponde à condição $R = \omega / \upsilon_{cM} = 1$, em que $\upsilon_{cM}/N = 3.2 \times 10^{-7}$ cm^3/s constitui o valor máximo de υ_c/N (que ocorre para $u \simeq 12$ eV). No semi-plano superior (R>1) a FDE torna-se unicamente função de E/ω (regime HF), enquanto que no inferior (R<1) será unicamente função de E/N, exceptuando obviamente uma zona de transição entre as duas regiões. A condição $\omega > \upsilon_{cM}$ poderá ser, todavia, demasiado severa para o estabelecimento do regime HF, e poder-se-ia substituir, por exemplo, por uma condição do tipo $\omega > \bar{\upsilon}_c$, em que $\bar{\upsilon}_c$ é definido pela Eq. (33). A curva B corresponde a $\omega = 10\ \bar{\upsilon}_c$; o regime HF estará certamente estabelecido no semi-plano acima desta curva. Finalmente, a linha vertical corresponde ao valor de E/ω para o qual a perda de energia por colisão elástica e inelástica é precisamente igual, em regime HF; para valores de E/ω inferiores (zona à esquerda) a perda elástica predomina e o corpo da distribuição torna-se Maxwelliano com temperatura $T_e = (e\ M/3\ m^2)\ (E/\omega)^2$[16, 24]; indicam-se no diagrama os valores de T_e em correspondência com os de E/ω, nesta situação.

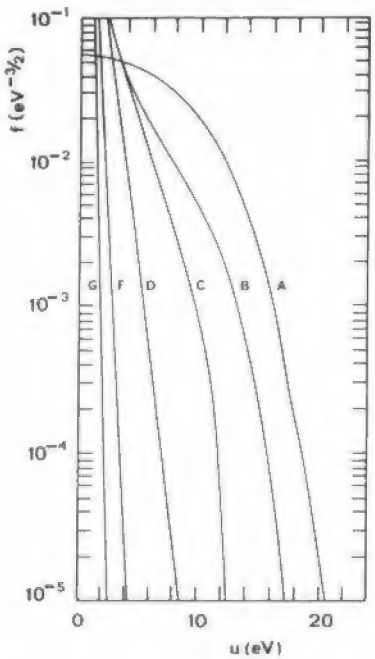


Fig. 8 Função de distribuição de energia dos electrões no argon para $E/N = 6{,}5 \times 10^{-16}$ V cm^2 e $R = \omega/\upsilon_{cM} = 0$ (A); 1,25 (B); 9,35 (C); 15,6 (D); 23,4 (F); 31,2 (G)

A forma da FDE modifica-se substancialmente à medida que se passa da situação $\omega \ll \upsilon_c$ para a de $\omega \gg \upsilon_c$. A Fig. 8 mostra a FDE no argon para um valor constante $E/N = 6{,}5 \times 10^{-16}$ Vcm^2 e vários valores de $R = \omega/\upsilon_{cM}$, desde zero (DC) a 31,2 (HF). À medida que R aumenta o ponto de funcionamento no diagrama da Fig. 7 desloca-se horizontalmente da direita para a esquerda. desde a região DC para a região HF. O corpo da FDE torna-se uma função rapidamente decrescente

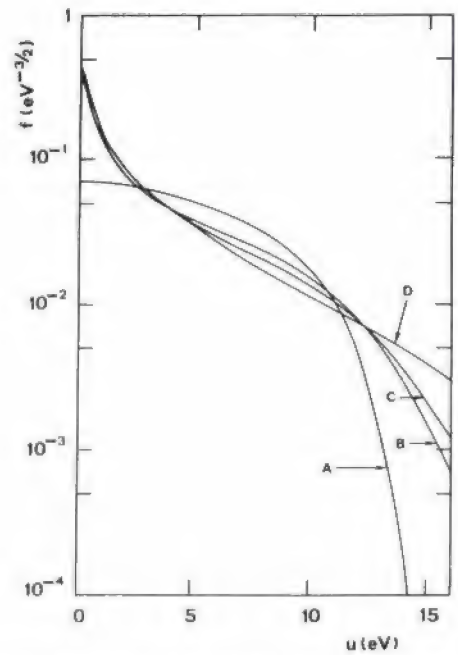


Fig. 9 Função de distribuição de energia dos electrões no argon para uma mesma energia média de 3,5 eV e os seguintes valores de $\omega/\upsilon_{ce} = 0$ (A); 0,5 (B); 0,8 (C); $\gg 1$ (D)

da energia, o que se traduz por um aumento significativo da fracção de electrões com energias elevadas (cauda da FDE). A razão física disso é fácil de entender: a potência transferida $\upsilon_c\ u_c$ é máxima quando $\upsilon_c(u) = \omega$, logo, sendo $\upsilon_c(u)$ uma função crescente até $u \simeq 12$ eV, a transferência de potência em regime HF é maior para electrões com energias da ordem da dezena de electrões-volt, o que enriquece substancialmente a cauda da distribuição.

Este enriquecimento relativo da cauda é bem demonstrado na Fig. 9 em que se mostram várias distribuições com a mesma energia média, nos casos DC, HF, e intermédio. Uma consequência importante deste facto é que os processos de excitação de limiar mais elevado e a ionização se tornam mais eficazes em regime HF, como se ilustra na Fig. 10. Nesta figura mostra-se a fracção de potência perdida pelos diversos canais colisionais nos casos DC e HF, em função da potência total dissipada por colisões, em média por electrão, para uma densidade unitária do gás, Θ/N*; as curvas correspondentes ao caso $\omega \gg \upsilon_c$ estão transladadas para menores valores de Θ/N em relação às do caso $\omega \ll \upsilon_c$.

Como se pode verificar, para um dado Θ/N a fracção de potência gasta na ionização é maior no caso HF que no DC; expressando isto noutra forma, para manter uma descarga compensando as perdas de electrões

* Note-se que

$$\frac{\Theta}{N} = (e\ n_e)^{-1}\ \mathrm{Re}\left(\frac{\sigma}{N}\right) E^2 = \frac{2\ m}{M} \left\langle u \frac{\upsilon_c}{N} \right\rangle + \sum_j V_j \left\langle \frac{\upsilon_j}{N} \right\rangle.$$

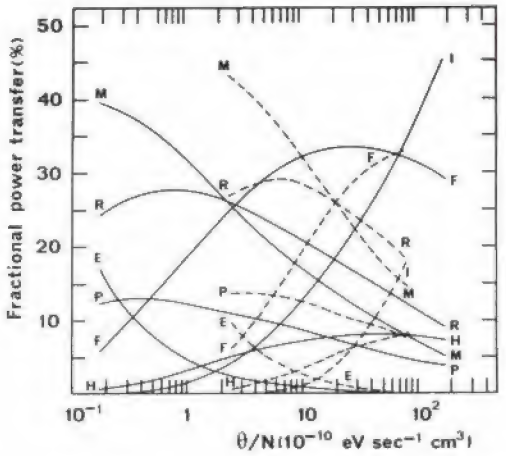


Fig. 10 Perdas de energia relativa dos electrões no argon pelos diferentes canais colisionais em funções da potência média absorvida por electrão para densidade unitária do gás, nos limites $\omega >> \upsilon_{ce}$ (a cheio) e $\omega << \upsilon_{ce}$ (a tracejado): E — colisões elásticas; M — $^3p_0 + {}^3p_2$; R — $^1p_1$; P — $^3p_1$; F — transições proibidas; H — estados permitidos elevados; I — ionização

pela ionização em volume, é necessária uma menor potência média absorvida por electrão no caso HF do que no caso DC[30]. Para uma mesma potência absorvida por unidade de volume o plasma HF terá, portanto, maior densidade que o plasma DC, assim como uma maior população de estados de limiar de excitação elevado. Estas conclusões concordam bem com diversos factos experimentais referidos na literatura[31].

Para o balanço de criação-perda dos pares electrão-ião (problema que, como referimos, é fundamental no estudo do equilíbrio do plasma) é necessário conhecer o coeficiente de ionização por colisão electrónica $C_i = <\sigma_i\ v>$, em que σ_i é a secção eficaz de ionização. Em princípio C_i, já que envolve uma média sobre a FDE, deverá exibir um tipo de dependência em relação a E/N e E/ω (ou E/N e υ_{ce}/ω) como o da própria FDE. No argon, contudo, Ferreira e Loureiro[30] mostraram que é possível definir um campo efectivo para a ionização da forma (note-se a analogia com o campo efectivo definido no caso $\upsilon_c(u) = \text{const.}$)

$$E_e^i = E \frac{\upsilon_{ce}}{(\upsilon_{ce}^2 + \omega^2)^{1/2}} \qquad (42)$$

com $\upsilon_{ce}/N = 2 \times 10^{-7}\,cm^3/s$, o qual permite expressar C_i como função da única variável E_e^i/N, para qualquer frequência (isto é, qualquer valor de ω/υ_{ce}). A Fig. 11 mostra o coeficiente de ionização no argon, obtido a partir da resolução da equação de Boltzmann[30].

3.3 Balanço global de energia

O balanço global de energia duma descarga HF pode escrever-se na forma[14, 32]

$$P_I\ A(\bar{n}_e) = \bar{\Theta}\ V_p\ \bar{n}_e,$$

em que P_I é a potência HF incidente no plasma, $A(\bar{n}_e)$ é a fracção dessa potência que é absorvida no volume do plasma V_p, $\bar{n}_e$ é a densidade média em V_p, e Θ representa a potência média perdida por electrão em colisões de todas as espécies no volume V_p. Em princípio Θ depende explicitamente de quantidades como a pressão do gás, as dimensões do recipiente, a densidade média $\bar{n}_e$, a frequência de funcionamento e ainda, em princípio, das distribuições espaciais da densidade do plasma e do campo HF no interior do plasma. Esta quantidade desempenha um papel fundamental na teoria dos plasmas HF pelo que, no seguimento, a analisaremos em pormenor no caso particular dos plasmas de 0. S.

O termo no primeiro membro da Eq. (42) representa, por assim dizer, o comportamento electrodinâmico do aplicador quando tem por carga o próprio plasma; deste ponto de vista, o plasma é fenomenologicamente representado como um dieléctrico cuja permitividade, para uma dada frequência, é unicamente determinada pela sua densidade média e pela frequência de colisão. A fracção de potência incidente que é absorvida, $A(\bar{n}_e)$, depende, naturalmente, do tipo de aplicador e de instala-

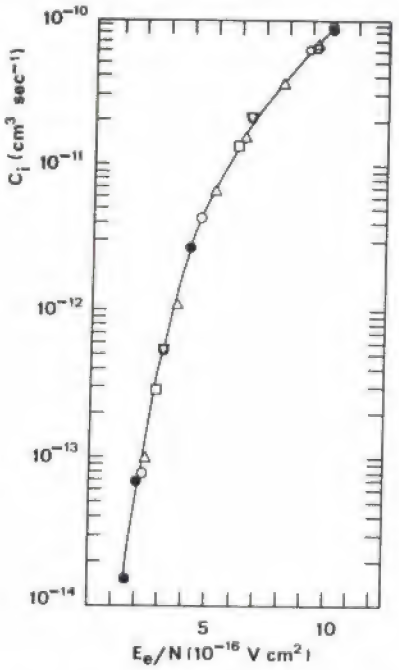


Fig. 11 Coeficiente de ionização no argon em função do campo efectivo reduzido para vários valores de ω/υ_{ce} entre zero e infinito (pontos). A curva representa a melhor interpolação destes pontos

ção. Assim, por exemplo, no caso de plasmas produzidos dentro duma cavidade ressonante, A $(\bar{n}_e)$ tem a forma duma curva de ressonância cuja amplitude depende do acoplamento entre o excitador e a cavidade.

No caso dos plasmas de O. S., ou de onda progressiva em geral, A $(\bar{n}_e)$ pode ser resolvido espacialmente.

Com efeito, designando por P (z) o fluxo total de potência através do plano z = const., a potência absorvida numa fatia de plasma entre z e z + dz pode escrever-se

$$P(z)\ dA(\bar{n}_e) = 2\alpha(\bar{n}_e)\ P(z)\ dz, \qquad (43)$$

em que α é o coeficiente de atenuação, anteriormente definido pela Eq. (36). Tendo em conta a Eq. (42), o balanço global de potência na fatia de plasma poderá, pois, escrever-se na forma

$$2\alpha(\bar{n}_e)\ P(z) = \bar{\Theta}\,\bar{n}_e\,\pi\ a^2. \qquad (44)$$

Mas, a potência absorvida por unidade de volume pelos electrões é dada por $dP_{abs.}/dV = R_e(\sigma)\ E^2$, o que se pode ainda escrever, no caso de υ_c ser independente de u,

$$\frac{dP_{abs}}{dV} = \frac{e^2 n_e}{m}\ \frac{E_e^2}{\upsilon_c}, \qquad (45)$$

sendo E_e o campo efectivo[40]. Integrando a Eq. (45) na secção transversal e combinando com (44) obtém-se facilmente

$$\bar{\Theta} = \frac{e^2}{m\,\upsilon_c} \int_0^1 \frac{n_e}{\bar{n}_e}\ E_e^2\ 2x\ dx. \qquad (46)$$

Convém escrever esta equação duma forma um pouco diferente que será útil para pôr em relevo a existência de relações de semelhança. Assim, escrevê-la-emos na forma

$$\frac{\Theta}{N} = \frac{e^2}{m\,(\upsilon_c/N)} \left(\frac{E_{eo}}{N}\right)^2 \int_0^1 \frac{n_e}{\bar{n}_e}\ E^2\ 2x\ dx, \qquad (47)$$

em que E_{eo} é o campo efectivo no eixo e $E = E_e/E_{eo}$.

Verifica-se, assim, que $\bar{\Theta}/N$ é uma função explícita de E_{eo}/N, o qual, como veremos, constitui um valor próprio do problema de valores na fronteira que resulta do equilíbrio criação-perda dos pares electrão-ião. Por outro lado, Θ/N depende, através do termo integral, das distribuições espaciais da densidade e da intensidade do campo no plasma. Estas distribuições espaciais, tal como o valor de E_{eo}/N, podem ser consistentemente determinados a partir da formulação do equilíbrio criação-perda acima referido.

Quando $\upsilon_c(u) \neq$ const. os raciocínios são idênticos mas as expressões são mais complicadas, por não se poder fazer pleno uso do conceito de campo efectivo. Só não será assim no caso $\upsilon_c(u) << \omega$, ou seja, para fraca atenuação, vindo neste caso [ver Eqs. (32)-(34)]

$$\frac{\bar{\Theta}}{N} = \frac{e^2}{m} \left(\frac{E_o}{\omega}\right)^2 \int_0^1 \frac{n_e}{\bar{n}_e}\ E^2\ \frac{\upsilon_c}{N}\ 2x\ dx, \qquad (48)$$

em que E_o designa agora o valor eficaz da intensidade do campo no eixo e $E = E/E_o$.

3.4 Equilíbrio de ionização-perda. Campo de manutenção

Num plasma quase-neutro (comprimento de Debye, λ_D, muito menor que qualquer dimensão do recipiente) o transporte dos electrões e dos iões para a parede pode ser descrito pelo modelo hidrodinâmico de dois momentos, isto é, usando as equações de continuidade e de transporte de quantidade de movimento para os electrões e os iões[33]. Estas podem escrever-se

$$\bar{\nabla} \cdot (n\bar{v}) = \upsilon_i\ n \qquad (49)$$

$$\bar{v} = -\mu_e\ \bar{E}_A - \frac{1}{n}\ \bar{\nabla}\ (D_e n) \qquad (50)$$

$$(\bar{v} \cdot \bar{\nabla})\bar{v} + \upsilon_i \bar{v} = \frac{e}{M}\ \bar{E}_A - \frac{K\,T_i}{M}\ \frac{1}{n}\ \bar{\nabla}\, n - \upsilon_{in}\bar{v}, \qquad (51)$$

em que $n = n_e = n_i$ é a densidade do plasma, $\bar{v}$ é a velocidade média de transporte comum às duas espécies (transporte ambipolar), E_A é o campo de carga do espaço, μ_e e D_e são a mobilidade e o coeficiente de difusão dos electrões, $\upsilon_i = N\,C_i$ é a frequência de ionização, T_i é a temperatura dos iões, e υ_{in} é a frequência de colisão para a transferência de quantidade de movimento ião-neutro incluindo os efeitos da transferência de carga simétrica e das colisões elásticas.

Este sistema de equações permite fazer surgir naturalmente a fronteira plasma-bainha de carga de espaço (junto da parede)[33], aparecendo esta fronteira como os pontos em que a velocidade ambipolar v se torna igual à velocidade isotérmica iónica do som $[K(T_e + T_i)/M]^{1/2}$ (o que satisfaz à condição de Bohm).

Note-se que T_e representa, aqui, a temperatura dos electrões quando a FDE é Maxwelliana; quando assim não acontece T_e deve ser substituída pela energia característica $u_\kappa = D_e/\mu_e$[34].

Como a bainha de carga de espaço tem espessura desprezável ($\lambda_D << \Lambda$, sendo Λ o comprimento de difusão característico do recipiente) pode-se admitir que a condição $v = [K(T_e + T_i)/M]^{1/2}$ ocorre na parede; desta forma, para uma dada geometria do recipiente, as Eqs. (49)-(51) constituem um problema de valores na fronteira.

No caso dum plasma isotérmico (T_e espacialmente constante), por exemplo, fixando a densidade do gás N e supondo T_i conhecido, só existe solução para um valor de T_e bem determinado (o qual determina o valor do coeficiente de ionização). A solução do problema pode então expressar-se sob a forma usual duma característica de descarga de T_e em função de $N\Lambda$. Contudo, uma tal característica carece de sentido quando os electrões não são Maxwellianos: neste caso, como o coeficiente de ionização C_i e a energia característica u_κ são funções do campo reduzido (E/N no caso d.c. ou $\omega \ll \nu_c$; E/ω no caso $\omega \gg \nu_c$; E_e/N no caso de $\nu_c(u) =$ $=$ const., para qualquer ω), as Eqs. (49)-(51) determinam o campo reduzido em função de $N\Lambda$.

Verifica-se, assim, que o campo eléctrico no interior do plasma não tem um valor arbitrário, mas sim, um valor bem determinado (chamado campo de manutenção), quando se fixam as condições de funcionamento (N, ω, geometria, etc.).

Esse valor é precisamente aquele para o qual a frequência de ionização compensa exactamente a taxa de perda dos electrões e dos iões por transporte para a parede, onde se vão recombinar.

Se o plasma não for isotérmico, o que acontece nomeadamente no caso dos plasmas de 0. S. visto que o campo não é homogéneo no interior do plasma, o problema é um pouco mais complicado porque ν_i e u_κ variam espacialmente. Essa variação é determinada pela própria variação do campo, a qual, por sua vez, depende da distribuição espacial da densidade do plasma [ver, por exemplo, as Eqs. (23) e (24)]. Para tratar o problema duma forma consistente é, por isso, necessário resolver simultaneamente as Eqs. (23), (24) e (44)-(51), após ter

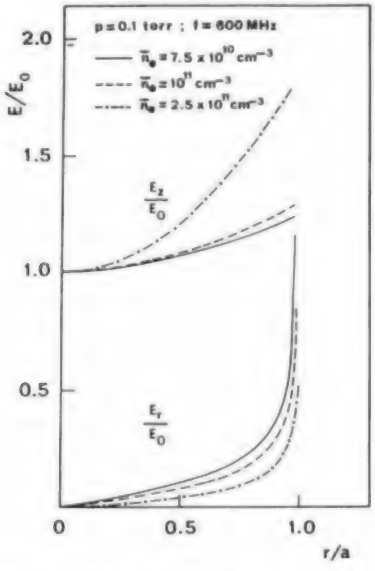


Fig. 12 Distribuição radial de densidade electrónica num plasma de argon mantido por uma 0. S. TM, para p = 0.1 Torr, f = 600 MHz e vários valores da densidade média

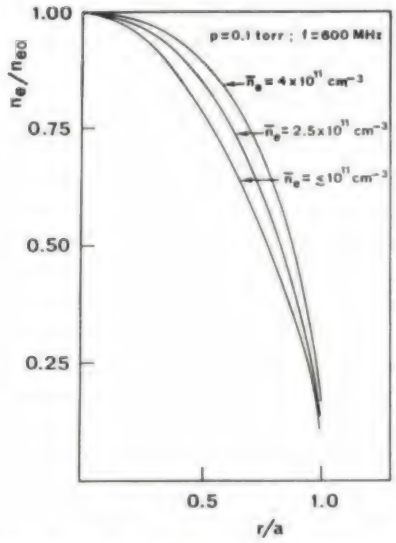


Fig. 13 Distribuições radiais das intensidades das componentes axial e radial do campo eléctrico duma O. S. TM propagando-se num plasma de argon inomogéneo, para diversos valores da densidade electrónica média

resolvido a própria equação de Boltzmann que analisámos no parágrafo 3.2. Note-se, no entanto, que a natureza do problema de valores na fronteira se mantém; neste caso, é agora o valor do campo reduzido no eixo do plasma que constitui um valor próprio do problema. Este campo de manutenção no eixo depende agora não só de N, ω, Λ, como no caso isotérmico, mas também da própria densidade média do plasma $\bar{n}_e$ (note-se que as Eqs. (23) e (24) mostram que a distribuição espacial do campo depende não só da distribuição espacial relativa da densidade mas também do seu valor absoluto, o qual pode ser medido, por exemplo, pelo seu valor no eixo ou pela média $\bar{n}_e$).

Nas figuras 12 a 15 mostram-se resultados obtidos anteriormente pelo autor[35], através desta formulação consistente, para o caso dum plasma de argon mantido num tubo de pyrex ($\epsilon_g = 4.52$) cilíndrico por uma O. S. TM fracamente atenuada. À medida que aumenta a densidade média do plasma, o campo total $(E_z^2 + E_r^2)^{1/2}$ concentra-se cada vez mais junto da parede, o que aumenta consideravelmente os valores das frequências de ionização e de excitação por colisão electrónica nesta região em relação aos valores no eixo. Uma maior taxa de ionização na periferia aumenta os gradientes de difusão nesta região ao mesmo tempo que os diminui na zona central, o que é bem evidenciado pelos perfis de densidade obtidos.

Como se vê na Fig. 14 a inomogeneidade radial do plasma afecta consideravelmente a distribuição espacial do campo em relação à que se observaria se o plasma fosse homogéneo. Note-se ainda que a componente radial do campo é desprezável face à axial para $\bar{n}_e$ bastante acima da densidade do corte ($n_{ec} = 2{,}47 \times 10^{10}$ cm^{-3}, para f = 600 MHz e $\epsilon_g = 4{,}52$).

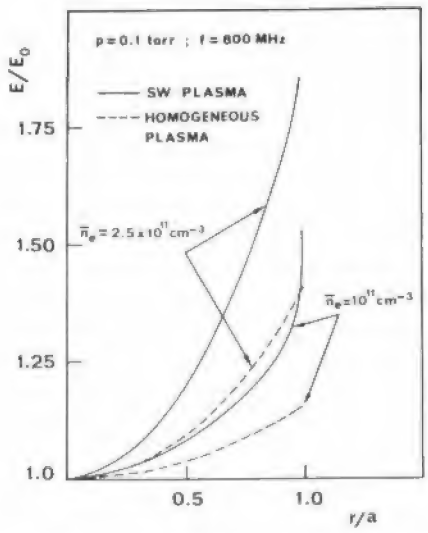


Fig. 14 Distribuições radiais da intensidade do campo eléctrico duma O. S. TM num plasma inomogéneo (a cheio) e num plasma homogéneo (a tracejado), para a mesma densidade electrónica média

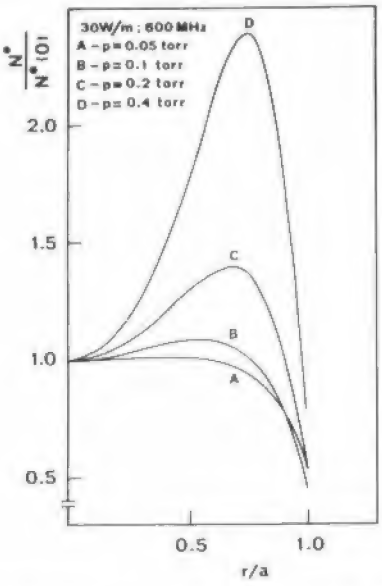


Fig. 15 Distribuições radiais da densidade de átomos de argon excitados na configuração $3p^5 6d$ num plasma mantido por uma O. S. TM, para uma mesma potência absorvida por unidade de comprimento e para várias pressões

A Fig. 15 mostra perfis radiais calculados de átomos excitados na configuração $3p^5\,6d$ para várias condições de funcionamento. A forma destes perfis pode ser facilmente compreendida atendendo a que a população destes estados é, localmente, proporcional ao produto $n_e C_j$, sendo C_j o coeficiente de excitação electrónica correspondente. A densidade n_e decresce com o raio, mas C_j aumenta e tanto mais rapidamente quanto maior for $\bar{n}_e$; o produto $n_e C_j$ poderá pois passar por um máximo para certa posição radial, o qual será tanto mais acentuado quanto maior for $\bar{n}_e$. É precisamente o que ilustra a Fig. 15, já que o aumento de pressão, mantendo a potência absorvida por unidade de comprimento, se traduz por um aumento de $\bar{n}_e$. Estes perfis calculados estão inteiramente de acordo com observações experimentais usando técnicas de espectroscopia de emissão[36, 37].

3.5 Leis de Semelhança

O plasma de O. S. fracamente atenuada ($\omega >> \upsilon_c$) obedece a leis de semelhança simples que permitem a extensão de resultados obtidos numa dada situação particular a outras situações experimentais[38]. Para prová-lo, convém escrever as equações básicas numa forma apropriada.

a) Equações de continuidade e de transporte de quantidade de movimento:

$$\frac{1}{x}\frac{d}{dx}(n\,v\,x) = (Na)\,C_i\,n \tag{52}$$

$$v = \frac{D_e N}{Na}\left(\frac{e}{K T_e}\frac{d\phi}{dx} - \frac{1}{n}\frac{dn}{dx}\right) \tag{53}$$

$$v\frac{dv}{dx} = -\frac{e}{M}\frac{d\phi}{dx} - \frac{K T_i}{M}\frac{1}{n}\frac{dn}{dx} - (Na)\left(C_i + \frac{\upsilon_{in}}{N}\right) v. \tag{54}$$

b) Equações do campo eléctrico da O. S. no plasma:

$$\frac{d^2 E_z}{dx^2} + \left(\frac{1}{x} + \frac{\beta^2}{\Gamma^2}\frac{1}{n}\frac{dn}{dx}\right)\frac{dE_z}{dx} - (\Gamma a)^2 E_z = 0 \tag{55}$$

$$E_r = \frac{\beta a}{(\Gamma a)^2}\frac{d E_z}{dx}. \tag{56}$$

c) Balanço local de energia dos electrões:

$$\frac{\Theta}{N} = \frac{e}{m}\frac{\bar{\upsilon}_c}{N}\left(\frac{E}{\omega}\right)^2 = \frac{2m}{M}\left\langle u\frac{\upsilon_c}{N}\right\rangle + \sum_j V_j C_j \tag{57}$$

Nestas equações $x = r/a$ é a coordenada radial normalizada, $E_z = E_z/E_o$ e $E_r = E_r/E_o$ são as intensidades das componentes do campo normalizadas em relação à intensidade no eixo, $\phi(x)$ é o potencial do plasma, υ_c e Γ são as grandezas definidas nas Eqs. (33) e (28), respectivamente.

Como já vimos, quando $\omega_{pe}^2 >> \omega^2$ verifica-se que $\Gamma \simeq \bar{\omega}_{pe} a/c >> \beta$. Na prática este limite é logo atingido quando ω_{pe} excede um pouco o valor de corte

$\omega_{pc} = \omega(1 + \epsilon_g)^{1/2}$, pelo que esta aproximação é válida em quase toda a coluna de plasma (exceptuando apenas a região em que esta termina e, eventualmente, a região periférica junto da parede do tubo).

Neste caso as Eqs. (55) e (56) podem escrever-se aproximadamente

$$\frac{d^2 E_z}{dx^2} + \frac{1}{x}\frac{d E_z}{dx} - \left(\frac{\bar{\omega}_{pe} a}{c}\right)^2 \frac{n}{\bar{n}} E_z = 0 \qquad (58)$$

$$E_r << E_z. \qquad (59)$$

Desta forma verifica-se que apenas existem dois parâmetros independentes no problema: Na e $\bar{\omega}_{pe} a/c$. Tal facto significa que as distribuições espaciais de densidade $n_e(x)/\bar{n}_e$, de intensidade do campo $E(x)$ e de temperatura $T_e(x)$, assim como a intensidade do campo reduzido no eixo E_o/ω, são unicamente funções daqueles dois parâmetros.

Sendo assim, imediatamente se conclui que a frequência média de colisão para densidade de gás unitária, $\bar{\upsilon}_c/N$, e a potência média absorvida por electrão numa fatia dz de plasma, para densidade de gás unitária, $\bar{\Theta}/N$ [ver Eq. (48)], são unicamente funções de Na e de $\bar{\omega}_{pe} a/c$:

$$\frac{\bar{\upsilon}_c}{N} = f_\upsilon \left(Na, \frac{\bar{\omega}_{pe} a}{c}\right) \qquad (60)$$

$$\frac{\bar{\Theta}}{N} = f_\Theta \left(Na, \frac{\bar{\omega}_{pe} a}{c}\right). \qquad (61)$$

Estas conclusões constituem leis de semelhança para este tipo de descarga. Verifica-se que estão plenamente de acordo com os resultados experimentais, como a seguir se mostra para o caso do argon.

A Fig. 16 mostra uma comparação dos valores médios radiais de $\bar{\upsilon}_c/p$ teóricos com dados obtidos de várias experiências[17, 23] (p é a pressão para uma temperatura de 300 k, isto é, p = 1 Torr corresponde a $N = 3{,}22 \times 10^{16}\,cm^{-3}$). O acordo entre teoria e experiência é satisfatório, o que indica, em particular, que a onda é principalmente atenuada pelas colisões electrónicas à medida que se propaga.

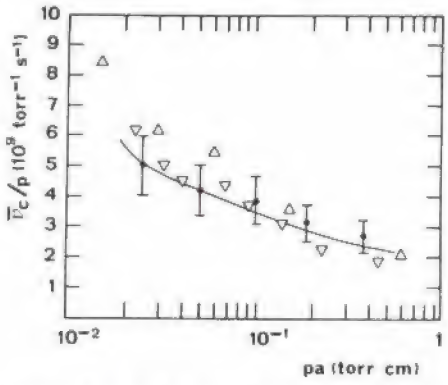


Fig. 16 Frequência de colisão efectiva a pressão unitária num plasma de argon mantido por uma O. S. TM em função do produto pa. Curva a cheio-teoria; pontos - resultados experimentais das refs.[17, 23]

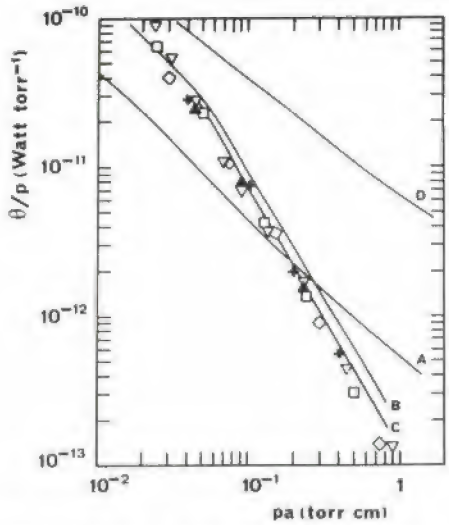


Fig. 17 Potência média absorvida por electrão a pressão unitária requerida para manter uma coluna cilíndrica de plasma de argon em função do produto pa, nas seguintes situações: i) Plasma isotérmico: $\omega >> \upsilon_{ce}$, FDE não Maxwelliana (A); $\omega >> \upsilon_{ce}$, FDE Maxwelliana (B); $\omega << \upsilon_{ce}$, FDE não Maxwelliana (D); ii) Plasma mantido por O. S. TM, para $\omega >> \upsilon_{ce}$: teoria com FDE Maxwelliana (C); resultados experimentais de Dervisevic et al[39] (pontos)

A Fig. 17 mostra, para fins de comparação, valores calculados de $\bar{\Theta}/p$ em função de pa para diversas situações (casos DC e HF supondo o plasma isotérmico, plasma de O. S. com electrões Maxwellianos e plasma isotérmico com electrões Maxwellianos). No caso do plasma de O. S. seria de esperar que $\bar{\Theta}/p$ dependesse também de $\bar{\omega}_{pe} a/c$, mas esta dependência é, de facto, tão ligeira que não se evidencia nos resultados da figura. Verifica-se que os cálculos de $\bar{\Theta}/p$ vs. pa no caso do plasma de O. S. com electrões Maxwellianos estão inteiramente de acordo com os resultados experimentais obtidos por Dervisevic et al[39], autores que investigaram plasmas de O. S. mantidos em capilares, usando várias frequências na banda das micro-ondas. (Note-se que nestes plasmas $\bar{n}_e$ é relativamente elevado, razão por que a FDE deve ser aproximadamente uma Maxwelliana).

Finalmente, a Fig. 18 mostra o campo efectivo reduzido no eixo $E_{eo}/N = (E_o/N)(\upsilon_{ce}/\omega)$ em função de Na, para dois valores de $\bar{\omega}_{pe} a/c$ admitindo que a FDE é Maxwelliana. Mostram-se, igualmente, para comparação, resultados obtidos para um plasma isotérmico nos casos DC e HF com a FDE obtida da equação de Boltzmann, assim como no caso HF com electrões Maxwellianos.

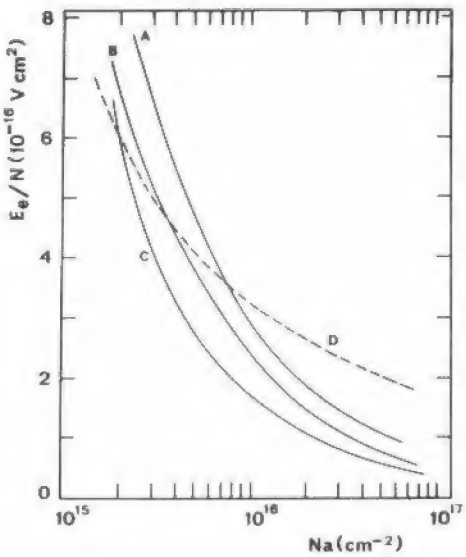


Fig. 18 Campo efectivo reduzido no eixo necessário para manter um plasma cilíndrico de argon em função do produto Na para $\omega >> \upsilon_{ce}$ e nas seguintes situações: i) Plasma isotérmico: FDE Maxwelliana (A) e FDE não Maxwelliana (D); ii) Plasma mantido por O. S. TM, com FDE Maxwelliana, para $\bar{\omega}_{pe}$ a/c = 0,75 (B) e $\bar{\omega}_{pe}$ a/c = 1,5 (C)

3.6 Propriedades axiais do plasma de onda de superfície

Nos parágrafos anteriores analisaram-se as propriedades radiais do plasma de O. S. e o equilíbrio de ionização-perda e de energia em cada fatia dz. É agora fácil descrever a evolução da densidade média $\bar{n}_e$ ao longo do eixo em função das condições de funcionamento. Com efeito, a partir da Eq. (44) obtém-se

$$\bar{n}_e = \frac{2\alpha\,(\bar{n}_e)\,P(z)}{\bar{\Theta}\,\pi a^2} \tag{62}$$

logo, derivando em ordem a z esta equação e combinando de novo com (44), virá

$$\frac{d\,\bar{n}_e}{dz} = -2\,\alpha\,\bar{n}_e\,(1 - \frac{\bar{n}_e}{\alpha}\,\frac{d\,\alpha}{d\,\bar{n}_e})^{-1}. \tag{63}$$

Desprezou-se aqui a dependência de $\bar{\Theta}$ em relação a $\bar{n}_e$ por ser, como vimos, muito fraca.

Usando as características de atenuação (Fig. 6) esta equação poderia ser facilmente integrada numericamente para obter $\bar{n}_e(z)$ em cada caso. A densidade no início da coluna $\bar{n}_e(0)$ (z = 0 corresponde à saída do surfatron) é simplesmente dada pela solução da equação

$$\bar{n}_e(o) = \frac{2\,\alpha_o\,P_i}{\bar{\Theta}\,\pi\,a^2}, \tag{64}$$

em que P_i é a potência total incidente em z = 0, $\bar{\Theta}/N$ é determinado pelo valor de Na e $\alpha_o = \alpha\,[\bar{n}_e(o)]$.

É, no entanto, útil evitar a integração numérica fazendo aproximações razoáveis que permitam obter expressões analíticas. Como se depreende da Fig. 6, log $(\alpha\omega/\bar{\upsilon}_c)$ varia quase linearmente com log $(\bar{\omega}_{pe}/\omega)$, excepto na vizinhança do corte (termo da coluna), pelo que usaremos a aproximação

$$\alpha\,\frac{\omega}{\bar{\upsilon}_c} = g\,(\omega,\,a)\,(\frac{\bar{\omega}_{pe}}{\omega})^{-h}, \tag{65}$$

em que g (ω, a) e h se obtêm facilmente, em cada caso, por interpolação dos resultados numéricos através da Eq. (65). A integração da Eq. (63) é então imediata vindo[38, 40]

$$\frac{\bar{n}_e\,(z)}{\bar{n}_e\,(o)} = (1 - \frac{h}{1 + (h/2)}\,\alpha_o\,z)^{2/h}. \tag{66}$$

Um caso particular interessante é h = 2; a atenuação α é então inversamente proporcional a $\bar{n}_e$ e o perfil axial de densidade é linear.

A Eq. (66) permite ainda obter uma expressão aproximada para o comprimento da coluna L, visto que $\bar{n}_e(L)$ deve aproximadamente corresponder à densidade de corte, $\bar{n}_{ec}$. Obtém-se facilmente

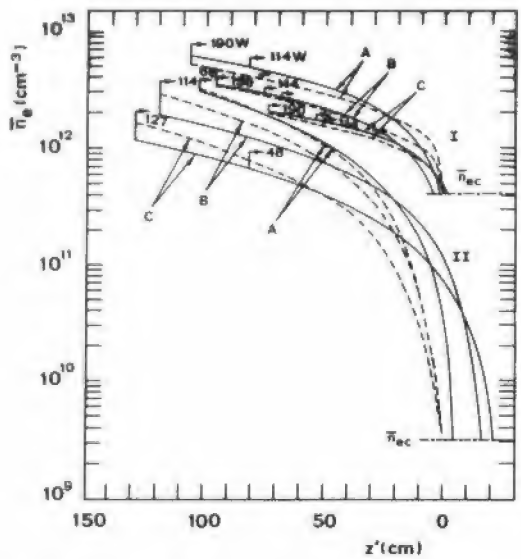


Fig. 19 Comparação das distribuições axiais calculadas (a cheio) e medidas (a tracejado) da densidade electrónica numa coluna de plasma de argon mantido por O. S. TM, para f = 2450 MHz (I) e 210 MHz (II) e para vários valores da potência incidente, em Watts, e da pressão, em Torrs (A - 0,5; B - 0,2; C - 0,1). Em abcissa representa-se a distância medida a partir do fim da coluna observado experimentalmente

$$L = \frac{1+(h/2)}{h} \frac{1}{\alpha_o} [1-(\frac{1+\epsilon_g}{Y_o})^{h/2}], \qquad (67)$$

em que $Y_o = (\overline{\omega}_{pe}(o)/\omega)^2$.

Apesar das aproximações que foram feitas este modelo simples descreve com rigor razoável a distribuição axial $n_e(z)$ e a influência dos vários parâmetros da descarga (N, ω, a, P_i, etc.) sobre esta distribuição e sobre o comprimento da coluna. A Fig. 19 mostra uma comparação dos resultados teóricos com os experimentais para diversos valores da frequência, pressão do gás e potência incidente (neste caso a = 0,45 cm, b = 0,60 cm e $\epsilon_g = 4.52$). Como era de esperar surgem algumas discrepâncias no fim da coluna, especialmente para a frequência mais baixa (210 MHz), em virtude das aproximações do modelo analítico. Tais discrepâncias poderiam, no entanto, ser eliminadas se se tivesse em conta o comportamento correcto da atenuação junto do corte.

4. CONCLUSÃO

O presente estado de desenvolvimento da teoria dos plasmas produzidos por ondas de superfície permite compreender os processos físicos fundamentais que entram em jogo e descrever com bastante rigor as propriedades radiais e axiais deste tipo de plasma em função dos parâmetros da descarga. Neste sentido, o estudo destes plasmas tem permitido progressos significativos na compreensão dos mecanismos essenciais que ocorrem em qualquer descarga de radiofrequência ou de micro-onda, abrindo desta forma o caminho para a formulação de modelos rigorosos e finos de plasmas deste tipo.

Esta possibilidade revela-se de grande interesse numa altura em que as descargas de radiofrequência e de micro-onda se começam a usar correntemente no mundo tecnológico para a gravura em microelectrónica, a deposição de filmes em superfícies, a excitação de Lasers, etc.

REFERÊNCIAS

1. L.D. LANDAU e E.M. LIFSHIFTZ, Electrodynamics of Continuous Media, Pergamon Press (Oxford 1960).
2. P.C. CLEMOW e J. ELGIN, J. Plasma Phys. *12*, 91 (1974).
3. A. SHIVAROVA e I. ZHELYAZKOV, Plasma Physics *20*, 1094 (1978).
4. A.W. TRIVELPIECE, Slow-Wave Propagation in Plasma Waveguides, California Inst. Techn., Pasadena, Technical Report 7 (May 1958).
5. M. MOISAN, A. SHIVAROVA e A.W. TRIVELPIECE, Plasma Physics *24*, 1331 (1982).
6. A.W. TRIVELPIECE e R.W. GOULD, J. Appl. Phys. *30*, 1784 (1959).
7. R.N. CARLILE, J. Appl. Phys. *35*, 1384 (1964).
8. C.C. PETERSON e J.F. DIENST, RCA Plasma Phys. Res. Eng. Tech. Papers (1966), pp. 18-25.
9. Y. AKAO e Y. IDA, J. Appl. Phys. *34*, 2119 (1963); ibid, *35*, 2565 (1964).
10. M. MOISAN, C. BEAUDRY e P. LEPRINCE, IEEE Trans. Plasma Sci. PS-3, 55 (1975).
11. Z. ZAKRZEWSKI, M. MOISAN, V.M.M. CLAUDE, C. BEAUDRY e P. LEPRINCE, Plasma Phys. *19*, 77 (1977).
12. M. MOISAN, Z. ZAKRZEWSKI e R. PANTEL, J. Phys. D: Appl. Phys. *12*, 219 (1979).
13. M. MOISAN, C.M. FERREIRA, Y. HAJLAOUI, D. HENRY, J. HUBERT, R. PANTEL, A. RICARD e Z. ZAKRZEWSKI, Rev. Phys. Appl. *17*, 707 (1982).
14. C.M. FERREIRA e M. MOISAN, "The Contribution of Surface Waves to the Modelling of RF and Microwave Discharges" in "Invited Lectures of the 2nd ICSW", World Science Publ. Co. (Singapore, 1986).
15. C.M. FERREIRA, "Gas Discharges Sustained by Surface Waves at Radio and Microwave Frequencies: Basic Processes and Modelling", in "Radiative Processes in Discharge Plasmas", NATO ASI Series, Plenum Press (New York, 1986).
16. W. P. Allis, Handb. of Phys. *21*, 383 (1956).
17. M. CHAKER, P. NGHIEM, E. BLOYET, P. LEPRINCE e J. MAREC, Rapport Interne L. P. 190, Université de Paris-Sud (Orsay, 1981), J. Phys. — Lett. *43*, L71 (1982).
18. P. NGHIEM, M. CHAKER, E. BLOYET, P. LEPRINCE e J. MAREC, J. Appl. Phys. *53*, 2920 (1982).
19. A.W. TRIVELPIECE, "Slow-Wave Propagation in Plasma Waveguides", San Francisco University Press (San Francisco, 1967).
20. D. ILIC, Int. J. Electronics *24*, 439 (1968).
21. O.M. GRADOV e L. STENFLO, Physics Reports *94*, 111 (1983).
22. G. BEKEFI, "Radiation Processes in Plasmas", Wiley (New York, 1966).
23. V.M.M. GLAUDE, M. MOISAN, R. PANTEL, P. LEPRINCE e J. MAREC, J. Appl. Phys. *51*, 5693 (1980).
24. C.M. FERREIRA e J. LOUREIRO, J. Phys. D: Appl. Phys. *16*, 2471 (1983).
25. I.B. BERNSTEIN e T. HOLSTEIN, Phys. Rev. *94*, 1475 (1954).
26. J.L. BLANCK, Phys. Fluids *11*, 1686 (1968).
27. T. HOLSTEIN, Phys. Rev. *70*, 367 (1946).
28. D.J. ROSE e S.C. BROWN, Phys. Rev. 98, 310 (1955).
29. S.C. BROWN. Hand. of Phys. *22*, 531 (1956).
30. C.M. FERREIRA e J. LOUREIRO, J. Phys. D: Appl. Phys. *17*, 1175 (1984).
31. A.P. MONTORNENKO. Opt. Spectrosc. *23*, 286 (1967).
32. M. Moisan e Z. ZAKRZEWSKI, "Gas Discharges Sustained by Surface Waves at Radio and Microwave Frequencies: Experiments" in "Radiative Processes in Discharge Plasmas", NATO ASI Series, Plenum Press (New York, 1986).
33. S.A. SELF e H.N. EWALD, Phys. Fluids *9*, 2486 (1966).
34. C.H. MULLER III e A.V. PHELPS, P. Appl. Phys. *51*, 6144 (1980).
35. C.M. FERREIRA, J. Phys. D: Appl. Phys. *14*, 1811 (1981).
36. M. MOISAN e A. RICARD, Can. J. Phys. *55*, 1010 (1977).
37. M. MOISAN, R. PANTEL, A. RICARD, V.M.M. GLAUDE, P. LEPRINCE e W.P. ALLIS, Rev. Phys. Appl. *15*, 1383 (1980).
38. C.M. FERREIRA, J. Phys. D: Appl. Phys. *16*, 1673 (1983).
39. E. DERVISEVIC, E. BLOYET, C. LAPORTE, P. LEPRINCE, J. MAREC, M. POUEY e S. SAADA, Proceed. XVI th Int. Conf. Phenom. Ionized Gases (Düsseldorf, RFA 1983) p. 468.
40. Z. ZAKRZEWSKI, J. Phys. D: Appl. Phys. *16*, 171 (1983).

USO DE MODELOS DE MARKOV NÃO OBSERVÁVEIS NO RECONHECIMENTO DE PALAVRAS ISOLADAS

António Joaquim Serralheiro*, José Manuel S. Tribolet**

RESUMO:

O reconhecimento de fala teve, em especial, a partir da última década um desenvolvimento notável. A finalidade do reconhecimento de fala é a de permitir a um utilizador dialogar com um sistema não humano.

Se bem que este objectivo ainda não tenha sido plenamente realizado, os resultados obtidos são encorajadores. O reconhecimento de palavras isoladas pode ser encarado como uma primeira etapa num dispositivo de reconhecimento de fala, uma vez que os objectivos a atingir são mais limitados.

Este trabalho descreve os sistemas geralmente utilizados: o Alinhamento Temporal Dinâmico, DTW, e Modelos de Markov Não Observáveis, HMM, com especial ênfase para este último.

ABSTRACT

In the last ten years, speech recognition has greatly improved. The final goal for a recognition system is the possibility of communication between man and machine.

Trully, such goal has not been achieved, but results are promising. Being less ambicious, isolated word recognition can be viewed as a first step towards the final system.

This paper describes the two most commonly used isolated word recognition systems: DTW (for Dynamic Time Warping) and, with special emphasis, HMM (for Hidden Markov Models).

* Assistente do IST.
** Professor Catedrático do IST.
Original recebido para publicação em 5/12/86.

0. PREÂMBULO

Este artigo decorre do trabalho de investigação efectuado pelos autores no âmbito do reconhecimento de fala. Nele se fará uma abordagem aos modernos sistemas de reconhecimento acústico global de palavras isoladas, detalhando-se os princípios teóricos e analíticos subjacentes ao uso de modelos de Markov não observáveis.

A investigação em Portugal neste domínio é extremamente recente. Com o objectivo de perspectivar este domínio de actividade, um dos autores (J. Tribolet) dedicou um semestre sabático nos Laboratórios Bell nos EUA à familiarização dos conceitos envolvidos em reconhecimento de fala e a investigar, em particular, o uso de algoritmos de programação dinâmica para alinhamento temporal de padrões (DTW — Dynamic Time Warping)[19], cuja tecnologia se encontra hoje amadurecida e já está comercialmente disponível.

O aparecimento mais recente de sistemas de reconhecimento baseados em modelos de Markov não observáveis (HMM — Hidden Markov Models) motivou o estudo e a familiarização de novos conceitos pelos autores, os quais são apresentados no presente trabalho.

Presentemente encontra-se em curso a tese de doutoramento de A. Serralheiro neste domínio, a decorrer no INESC, sob a orientação de J. Tribolet e em colaboração com os Laboratórios Bell.

1. INTRODUÇÃO

Neste trabalho far-se-á uma abordagem aos modernos sistemas de reconhecimento de fala, detalhando para o reconhecimento de palavras isoladas. Dos métodos ge-

ralmente utilizados, alinhamento temporal de padrões (DTW — Dynamic Time Warping) e modelos de Markov não observáveis (HMM — Hidden Markov Models), referiremos este último mais detalhadamente. Este trabalho principia por situar os sistemas de reconhecimento de fala no contexto do processamento digital, aborda algumas das dificuldades de implementação e seguidamente apresenta o modelo canónico do reconhecedor, após o que se pormenorizará cada um dos seus diferentes blocos operacionais.

O reconhecimento de fala exige duas fases distintas: a aprendizagem ou treino, e o teste ou classificação; na primeira o sistema é levado a criar as suas referências internas pelo que o(s) utilizador(es) deverá proferir a palavra ou palavras que deseje virem a ser reconhecidas; a duração de aprendizagem está relacionada com o número de palavras que se quer reconhecer e ainda com o desempenho que se pretende para o sistema reconhecedor. Uma vez concluída esta primeira fase, o sistema de reconhecimento encontra-se apto a executar a segunda fase que consta do reconhecimento propriamente dito. O reconhecimento de fala pode ser encarado ou como um problema clássico de reconhecimento de padrões, ou através de uma análise semântica do sinal acústico; esta segunda versão é por vezes acrescentada à primeira de modo a aumentar significativamente a qualidade do sistema de reconhecimento.

Este trabalho encontra-se estruturado do seguinte modo: no capítulo 2 são apresentadas algumas das motivações que impulsionaram o reconhecimento de fala, uma breve resenha dos trabalhos pioneiros sobre este assunto, alguns dos problemas a enfrentar quando se pretende implementar um sistema de reconhecimento. Referem-se os diferentes tipos de reconhecedores, e o modelo canónico de um sistema de reconhecimento de fala. Ainda sobre este modelo são descritos sumariamente alguns dos blocos constituintes. O capítulo 3 faz referência sucinta aos algoritmos de alinhamento temporal dinâmico. O capítulo 4 descreve o recurso aos HMM para o reconhecimento de palavras isoladas. O capítulo 5 é destinado às conclusões.

2. BREVES CONSIDERAÇÕES SOBRE RECONHECIMENTO DE FALA

Depois do aparecimento dos computadores digitais, começaram a ser desenvolvidos algoritmos de processamento de sinais, em particular de fala. Com este desenvolvimento, tornaram-se possíveis aplicações no campo de síntese e do reconhecimento de voz. A possibilidade de comunicar oralmente com um computador, ou com qualquer dispositivo, começou a ganhar viabilidade, sendo hoje comuns sistemas automáticos de resposta e de reconhecimento de palavras isoladas.

Tais sistemas permitem, por exemplo, uma maior rapidez de utilização de terminais de computador, uma vez que são eliminados os erros de escrita. Deste modo, possibilita-se o emprego das mãos para outras tarefas, o que é especialmente vantajoso no uso de determinado tipo de maquinaria, em que as ordens são dadas oralmente sem necessidade de deixar de se prestar atenção ao trabalho em curso. São ainda previsíveis outras aplicações, nomeadamente na prestação de serviços, tais como sistemas de reserva de passagens de avião, etc.

É interessante referir[1] que o primeiro reconhecedor de fala conhecido era um brinquedo — um cão que saltava quando o seu nome era proferido: "Radio Rex"; todavia, quando outros sons vagamente parecidos com "Rex" eram proferidos o dispositivo era activado. Provavelmente, a primeira tentativa séria de abordar o reconhecimento de fala foi efectuado em 1950 por Dreyfus-Graf em França; o primeiro sistema de reconhecimento de palavras isoladas mas que era optimizado para um dado utilizador foi construído nos laboratórios Bell em 1952 por Davis, Biddulph e Balashek permitindo um reconhecimento correcto de uma de dez palavras armazenadas em 97/100 das vezes. O primeiro trabalho que utilizava um computador digital data de 1959-1960, realizado por Denes e Mathews introduzindo o conceito do alinhamento temporal. Hoje, o reconhecimento de fala está intrinsecamente ligado ao processamento digital de sinais, efectuado quer em computadores de uso genérico "main frame general purpose", quer em processadores especializados, nomeadamente microprocessadores, tendo aparecido nos últimos anos integrados de elevada capacidade de processamento que permitem operar em tempo real.

Um sistema de reconhecimento de fala pode ser representado pelo seguinte diagrama de blocos:

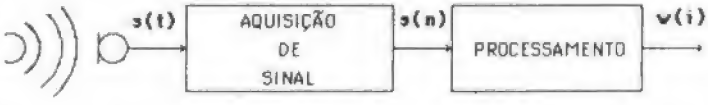


Fig. 1 Sistema de reconhecimento, diagrama de blocos

As ondas sonoras são captadas por um microfone, e o sinal resultante é primeiramente limitado em frequência, seguidamente sujeito a um processo de amostragem e quantificação, obtendo-se a sequência **s(n)**. Esta sequência é, em seguida, processada sendo **w(i)** o resultado deste processamento.

2.1 Dificuldades

Sendo a entrada do sistema um sinal acústico, deverão ser observados certos cuidados na sua recolha, referimo-nos à existência de ruído ambiente, por exemplo.

Assim, o microfone deverá ser de qualidade, permitindo a sua colocação o mais perto possível dos lábios, de modo a minimizar a captação do ruído, mas não tão próximo para não importunar o utilizador, impedindo a sua mobilidade e o seu conforto. No caso de se utilizar o micro-auscultador telefónico como entrada de comandos, é de esperar uma degração do nível de reconhecimento por parte do sistema, uma vez que à baixa qualidade destes transdutores se associam os diferentes tipos de ruído presentes nas linhas telefónicas comutadas.

Mencionámos alguns dos problemas mais óbvios relacionados com a recolha da informação oral; contudo, não podemos deixar de referir que o facto de se pretender alargar o uso destes dispositivos a um maior número de indivíduos, traz problemas adicionais relacionados com as variações de dicção existentes entre os diferentes utilizadores. Sendo de um modo geral um processo complexo, o reconhecimento de fala necessita, para ser implementado, de resposta a um determinado número de questões[15]:

1. *Que tipo de fala*? (contínua, palavras isoladas, etc.)
2. *Quantos oradores*? (um só, alguns, população em geral)
3. *Que tipo de oradores*? (cooperantes, casuais, treinados, homens, mulheres, etc.)
4. *Que tipo de ambiente acústico*? (sala sossegada, ruidosa, lugar público)
5. *Que tipo de equipamento*? (microfone de qualidade, micro-auscultador telefónico)
6. *Que necessidade de treino tem o sistema*? (algumas palavras, parágrafos, frases, etc.)
7. *Que treino para os utilizadores*? (adaptação natural, elaborada)
8. *Que dimensão para o vocabulário*? (50, 200, 1000, 10 000, palavras seleccionadas, ilimitado)
9. *Que tipo de linguagem*? (frases fixas, linguagem artificial, português natural, adaptado ao orador)
10. *Que serviço deve ser executado*? (respostas fixas, etc.)
11. *O que se sabe acerca do orador*? (motivação, etc.)
12. *Qual o grau de sofisticação do diálogo*? (resposta apenas, pedido de repetições)
13. *Que erros se podem tolerar*? (nenhuns [<0,1%], alguns [<10%], muitos [>20%])
14. *Qual a rapidez do sistema*? (tempo diferido, tempo proporcional à frase, tempo real)
15. *Qual o processamento envolvido*? (1 Mips, 10 Mips>10 Mips)
16. *Qual a memória necessária*? (1 M, 10 M, + [em bit/segundo de fala])
17. *Qual a sofistificação do software*? (programação simples, processos concorrentes, processamento paralelo, etc.)
18. *Qual o custo por segundo de fala*? (1, 10, 100, + em milhares de $US)
19. *Data de operacionalidade*?

2.2 Implementação

Os dispositivos de reconhecimento de fala podem agrupar-se, e sem detrimento de outras classificações, quanto ao tipo de operação em:

reconhecedores de palavras isoladas,
palavras ligadas,
fala contínua,

e quanto à sua relação com o utilizador em:

reconhecedores de fala independentes do utilizador,
dependentes do utilizador.

Mencionam-se, na literatura da especialidade, diferentes tipos de reconhecedores, que podemos agrupar em:

a) *reconhecedores de palavras isoladas*, em que cada palavra é pronunciada isoladamente das outras por meio de pausas (silêncios). Temos como exemplo **"Pára!"**.

b) *reconhecedores de palavras ligadas*, em que as fronteiras entre as palavras são menos nítidas, como por exemplo: **"Três mil e quarenta e um"** por oposição aos anteriores que obrigariam a **"Três-zero-quatro-um"**.

c) *busca de palavras chave*, em que se procura detectar a ocorrência de determinadas palavras, tais como: **"Ataque"**.

d) *reconhecedores de fala contínua*, mas num domínio restrito, em que se tem em vista um determinado objectivo: **"2 lugares para o voo..."**.

e) *idem*, mas qualquer que seja o contexto de aplicação: **"Qual o critério de escolha para adquirir...?"**.

f) *reconhecimento do utilizador*, em que se pretende que haja um acesso seguro ao sistema, por ex. aplicável no "cockpit" de um avião.

Um sistema de reconhecimento pode ser expresso através de um modelo canónico[3] que se indica na fig. 2b.

A entrada do sistema é o sinal acústico proveniente de uma fonação (geralmente uma palavra, uma frase), fig. 2a. A saída do modelo é a melhor ou as melhores estimativas da palavra ou das palavras proferidas. Por vezes a saída é ligada a um pós-processador como já se referiu.

O modelo apresentado tem sido largamente utilizado devido a várias razões, entre as quais poderemos referir que é invariante para vários vocabulários, conjuntos de parâmetros, algoritmos de semelhança entre padrões e regras de decisão.

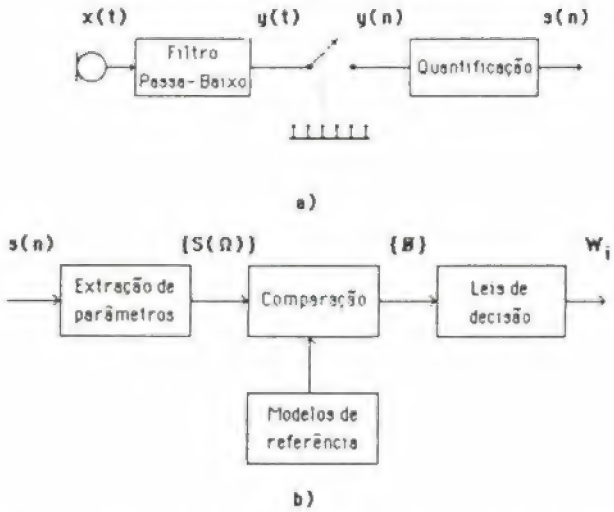


Fig. 2 a) aquisição do sinal. b) Modelo canónico para reconhecimento de fala

Por extracção de parâmetros entende-se um processo de redução de informação, em que a partir de um conjunto extenso de dados, nomeadamente amostras quantificadas da voz **s(n)** a um ritmo que poderá ser de 8000 amostras/segundo (p. ex.), se obtém outro conjunto mais pequeno de dados (valores típicos de redução 10 a 100) denominados parâmetros, e que descrevem algumas das propriedades do sinal acústico. No caso em questão, é vulgar o uso de:

a) espectro localizado (*transformada de Fourier*),
b) bancos de filtros,
c) coeficientes de predição linear, coeficientes Parcor (*algoritmos de predição linear*),
d) coeficiente cepstrais (*modelos homomórficos*).

Optar por qualquer destes, obedece a critérios que ponderam:

a) tempo de cálculo,
b) memória necessária,
c) facilidade de implementação,
d) qualidade a exigir ao sistema.

Nos sistemas comercializados são vulgarmente empregues a predição linear e a análise por banco de filtros, sendo esta última divulgada entre os primeiros sistemas por não necessitar de processamento digital para a extracção de parâmetros.

As regras de decisão permitem, uma vez conhecida uma lista de candidatos para a ou as palavras proferidas, obter o melhor candidato segundo critérios pré-estabelecidos[1, 3].

3. ALINHAMENTO TEMPORAL DINÂMICO

O alinhamento temporal dinâmico (DTW) consiste em comparar de algum modo um **padrão** de teste com padrões de referência. Por padrão entende-se uma colecção de conjuntos de parâmetros extraídos da fala. Por sua vez cada conjunto de parâmetros pode representar a característica espectral de um segmento de fala. Assim, poderemos dizer que na saída do primeiro bloco — extracção de parâmetros, teremos uma sucessão temporal de "espectros" que de algum modo caracterizam a palavra proferida. Como referências teremos, e para cada palavra do vocabulário, padrões que foram obtidos através de um processo de aprendizagem.

Consiste, este método, em determinar a "semelhança" entre os padrões de teste e de referência, o que é efectuado através de uma distância. Como a duração dos padrões varia, torna-se necessário um alinhamento temporal destes antes de se proceder à determinação da semelhança entre os mesmos. Consideremos dois padrões, um de teste **T (t)** e outro de referência **R (t)**; o objectivo é determinar a função de alinhamento **W (t)**, ou de "warping", que efectua a correspondência de **R** para **T**, baseando-se na minimização da distância entre os dois padrões:

$$\mathbf{D\ (T, R) = \min_{\{w\ (t)\}} \int d[t,\ W(t)]\ G\ [t,\ W(t),\ W'(t)]\ dt}$$

sendo **{W (t)}** o conjunto de todas as funções monotonicamente crescentes, com derivada contínua, **d [t, W (t)]** uma distância pontual de **R** a **T, G** é uma função de ponderação.

Seja **T = {T (1), T (2),..., T (NT)}** o conjunto dos padrões de teste e **R = {R (1), R (2),..., R (NR)}** o conjunto dos padrões de referência. O alinhamento temporal óptimo é a curva que relaciona o eixo temporal **m** do padrão de teste **T (m)** com o eixo dos tempos **n** do padrão de referência, na forma:

$$\mathbf{m = W(n)}$$

Como cada padrão é constituído por um conjunto contável de "espectros": tantos quantos os segmentos em que se partiu o sinal correspondente a cada palavra, a determinação da função de alinhamento consiste na determinação de um caminho numa rede. Esta determinação envolve a especificação de vários factores:

a) restrições aos pontos inicial e terminal,
b) restrições locais: "percursos possíveis",
c) restrições globais: limitação do espaço,
d) escolha dos eixos para os padrões,
e) medida das distâncias.

Não detalharemos os pontos anteriores por se encontrar fora do âmbito deste trabalho, veja-se[1, 3, 16].

3.1 Implementação do algoritmo de DTW

Para se implementar o algoritmo de alinhamento temporal dinâmico, utilizando um procedimento de progra-

mação dinâmica, consideram-se válidas as seguintes asserções:

a) um percurso globalmente óptimo também o é localmente,

b) o caminho óptimo para o ponto da rede **(n, m)** só depende dos valores de **n'**, **m'** tais que: $n' \leqslant n$, $m' \leqslant m$.

Define-se assim a relação recursiva de programação dinâmica; seja $D_A(n, m)$ a distância parcial acumulada desde **(1,1)** até **(n, m)**:

$$D_A(n, m) = d(T_n R_m) + \min_{q \leqslant m} [D_A(n-1, q)]$$

$D_A(n, m)$ pode ser definida recursivamente em termos de um ponto intermédio **n'**, **m'** $[n' \leqslant n, m' \leqslant m]$:

$$D_A(n, m) = \min_{(n', m')} [D_A(n', m') + \hat{d}((n', m'), (n, m))]$$

em que $\hat{d}$ é a distância ponderada de $(n', m')-(n, m)$ e **L** é o número de segmentos no percurso anterior, e $\tilde{W}(n)$ é uma função de ponderação:

$$\hat{d}[(n', m'), (n, m)] = \sum_{l=0}^{L-1} d[(n-l), (m-l)]\, \tilde{W}(n')$$

Para uma implementação eficiente de $D_A(n, m)$ é apenas necessário aplicar restrições à variação do percurso de (n', m') para (n, m) dentro da rede. Uma das restrições utilizadas é a de Itakura, em que os percursos possíveis de (n', m') para (n, m) são os indicados abaixo; note-se que o ramo assinalado com **x** corresponde a não permitir que o percurso seja horizontal durante mais do que um segmento.

Fig. 3 Percursos possíveis, restrição de ITAKURA.

$$D_A(n, m) = (\min\{D_A(n-1, m-2) + d(n, m),$$

$$D_A(n-1, m-1) + d(n, m), D_A(n-1, m)\, g(k) + d(n, m)\}$$

sendo $g(k) = \infty$ no ramo assinalado com **x**, caso contrário, $g(k) = 1$.

Para a implementação do algoritmo teremos:

a) início: $D_A(1, 1) = d(1, 1)\, \tilde{W}(1)$

b) recursividade: calcule $D_A(n, m)$ recursivamente para $1 \leqslant n \leqslant N$, $1 \leqslant m \leqslant M$

c) finalização:

$$\hat{D} = D_A(N, M) = \frac{1}{N(\tilde{W})}$$

sendo $N(\tilde{W})$ uma função de normalização.

3.2 Algumas considerações sobre o DTW

Para ilustrar a determinação de $D_A(n, m)$, atente-se no seguinte exemplo, em que temos para a palavra de teste FOUR pronunciada *FORR* e que como referência temos *FFOOR*, e admitamos que cada letra é um bloco, que os pontos terminais foram correctamente determinados, e que se obtinha a seguinte matriz de verossemelhança:

(Referência)				
R	9	6	2	2
O	8	1	8	6
O	8	1	8	6
F	2	3	7	7
F	1	7	7	3
	F	O	R	R
	(Teste)			

Fig. 4a) Matriz de verossemelhança.

da qual se obtém a matriz das distâncias acumuladas mínimas:

(Referência)				
R	28	11	7	9
O	19	5	13	18
O	11	4	12	18
F	3	9	15	22
F	1	8	15	20
	F	O	R	R
	(Teste)			

Fig. 4b) Matriz das distâncias acumuladas mínimas

Teríamos deste modo que a distância acumulada mínima entre as duas palavras é de 9. Procederíamos de igual modo para os restantes padrões de referência, obtendo várias distâncias. Estas seriam sujeitas a um processo de decisão que elegeria um ou mais candidatos para a palavra *FORR*, e se não ocorresse nenhum erro de reconhecimento, o sistema reconhecê-la-ia como *FOUR*.

O desempenho de sistemas de reconhecimento de palavras isoladas é bastante bom, havendo sistemas comercializados[20] em que a taxa de reconhecimento correcto atinge os 99%, para um só utilizador. O aparecimento de circuitos integrados dedicados ao reconhecimento por

DTW[21] vem permitir uma maior difusão dos sistemas de reconhecimento.

Quando a recolha do sinal acústico é efectuada num ambiente ruidoso, quando se trata do reconhecimento de palavras ligadas, ou ainda com alguns sons pode tornar-se difícil a localização exacta dos início e fim das palavras, e como tal podem ocorrer erros no reconhecimento, uma vez que se estão a atribuir distâncias elevadas aos primeiros segmentos dos padrões, e eventualmente podem ser logo excluídas as palavras correctas. veja-se, p. ex.[3] para uma abordagem mais completa.

4. OS MODELOS DE MARKOV NÃO-OBSERVÁVEIS E O RECONHECIMENTO DE PALAVRAS ISOLADAS

Os processos de reconhecimento acústico global têm como base modelos de produção de fala; através de observações efectuadas na saída do aparelho fonador tenta-se inferir qual a palavra ou conjunto de palavras proferidas, pertencentes a um dado vocabulário e que são o suporte da informação que se pretende transmitir.

De uma forma necessariamente simplista podemos supor que as palavras são proferidas a partir de um número finito de configurações do tracto vocal, e é o modo como essas configurações se sucedem ao longo do tempo que produzem os diferentes sons. Assim sendo, poderemos ainda supor que se num dado instante o tracto vocal está numa dada configuração há uma certa probabilidade de no instante de tempo seguinte se encontrar noutra configuração.

A ocorrência dos sons de uma dada palavra está, portanto, condicionada a um processo probabilístico. Somos deste modo levados a admitir um modelo probabilístico para o processo de produção da fala. A variação temporal das características espectrais do sinal de fala pode ser associada à lei probabilística de transição de estados de uma cadeia de Markov, enquanto que para cada instante de tempo, o espectro observado é condicionado a uma observação efectuada num dado estado ocupado. A ordem do modelo de Markov poderá ser qualquer, e as saídas dos seus estados podem ser um processo aleatório multivariável com densidades de probabilidade conjunta contínuas.

No que se segue restringir-nos-emos a modelos de primeira ordem e a observações que foram retiradas de um alfabeto discreto e finito de acordo com funções de distribuição discretas associadas a cada estado.

Começar-se-á por uma referência a conceitos fundamentais:

Definição: Um processo de Markov de 1.ª ordem é um processo estocástico cujo passado não tem influência no futuro se o presente é especificado[2]:

se $t_{n-1}<t_n$ **então**

$$P\{X(t_n) \leqslant x_n \mid X(t),\ t \leqslant t_{n-1}\} = P\{X(t_n) \leqslant x_n \mid X(t_{n-1})\}$$

donde se retira que se $t_1<t_2<...<t_n$ teremos:

$$P\{X(t_n) \leqslant x_n \mid X(t_{n-1}),..., X(t_1)\} = P\{X(t_n) \leqslant x_n \mid X(t_{n-1})\}$$

e para sistemas discretos no tempo:

$$P\{X_n \leqslant x_n \mid X_{n-1},..., X_1\} = P\{X_n \leqslant x_n \mid X_{n-1}\}$$

Definição: Uma função probabilística de uma cadeia de Markov não observável é um processo estocástico gerado por dois mecanismos interligados[12]:

— uma cadeia de Markov subjacente com um número contável de estados;
— um conjunto de funções aleatórias, cada uma das quais está associada a cada estado.

Assim, num dado instante **n** o processo é suposto estar num só estado e um símbolo é produzido de acordo com a função de distribuição de probabilidade que depende do estado actual. A cadeia de Markov subjacente muda de estado de acordo com a matriz probabilística de transição. O observador apenas vê como saídas (os símbolos anteriormente referidos) valores das funções aleatórias associadas a cada estado e não pode observar directamente os estados da cadeia de Markov, donde o nome de "não observável". Os valores das funções aleatórias correspondem, por exemplo, a conjuntos de coeficientes de LPC, associados a uma dada configuração do tracto vocal.

Suponha-se que uma cadeia de Markov tem N estados $q_1, q_2,..., q_N$ e que as observações são retiradas de um alfabeto **V** de **M** espectros $v_1, v_2,..., v_M$. A cadeia pode ser especificada em termos de um vector de distribuição de estados inicial $\pi' = (\pi_1, \pi_2,..., \pi_N)$ e uma matriz de transição de estados $A = [a_{i,j}]\ 1 \leqslant i, j \leqslant N$. O significado de π_i é o da probabilidade de nos encontrarmos num estado q_i num instante de tempo inicial, $t = 0$. Quanto a $a_{i,j}$ é a probabilidade de transitar (no instante seguinte) para o estado q_j estando no estado q_i:

$$a_{ij} = P(q_j \text{ em } t+1 \mid q_i \text{ em } t)$$

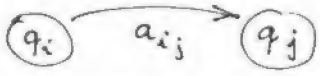


Fig. 5 Probabilidades de transição

O processo aleatório associado com os estados pode ser representado por uma matriz estocástica $B = [b_{j,k}]$ com $1 \leqslant j \leqslant N\ 1 \leqslant k \leqslant M$ sendo $b_{j,k}$ a probabilidade de observar o símbolo v_k sabendo que é q_j o estado actual:

$$b_{jk} = P(v_k \text{ em } t \mid q_j \text{ em } t)$$

Deste modo, o HMM é caracterizado pelo terno ordenado $(\pi, \mathbf{A}, \mathbf{B})$.

Para se poder utilizar os HMM para reconhecimento é necessário resolver dois problemas:

— estimação da sequência de observações, que será usada para a classificação de uma dada fonação,
— estimação dos parâmetros do modelo e que servirá para o treino deste.

Ambos os problemas provêm de uma sequência de observações $\mathbf{O}_1\ \mathbf{O}_2 \ldots\ \mathbf{O}_T$ em que cada $\mathbf{O}_t$ $(1 \leqslant t \leqslant T)$ é algum $v_k \in \mathbf{V}$, isto é uma observação correspondente a um símbolo do vocabulário.

Quanto à classificação, pretende-se reconhecer palavras de um vocabulário seleccionado de **W** palavras $\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_v$. Como entrada do sistema tem-se uma sequência de observações **O** e um conjunto de **v** modelos (um por cada palavra, pelo menos) $\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_v$. Tem de se calcular a probabilidade de a sequência ser gerada por cada um dos modelos: $\mathbf{P}_i = \mathbf{P}\{\mathbf{O}/\mathbf{M}_i\}$ $(1 \leqslant i \leqslant V)$. Se $\mathbf{P}_i \geqslant \mathbf{P}_j$ $1 \leqslant j \leqslant V$ a palavra de teste será reconhecida como a palavra $\mathbf{W}_i$; o mesmo é dizer que se reconhece como palavra pertencente ao vocabulário aquela que tiver a maior probabilidade de ter sido gerada pelo modelo **i**.

Quanto ao treino do sistema de reconhecimento, trata-se de determinar os modelos $\mathbf{M}_i = (\pi_i, \mathbf{A}_i, \mathbf{B}_i)$ $(1 \leqslant i \leqslant V)$, a partir de **K** referências de treino: $\mathbf{O}^{(1)}, \mathbf{O}^{(2)}, \ldots, \mathbf{O}^{(K)}$, sabendo que a observação $\mathbf{O}^{(i)}$ é proveniente de uma palavra (falada) $\mathbf{W}_i$ do vocabulário.

4.1 Classificação (teste)

Uma vez que o HMM lida com um número finito de símbolos, para se poder reconhecer fala recorrendo a este método torna-se necessário obter um conjunto numerável de símbolos. Assim, e se se estiver a empregar coeficientes de predição linear, estes terão de ser sujeitos a um processo de quantificação, por exemplo, a referida quantificação vectorial[10], [11]. Recorre-se à quantificação vectorial em vez de quantificação escalar, porque para uma mesma distorção aquela necessita de um menor número de índices.

Seja $\mathbf{O} = \mathbf{O}_1\ \mathbf{O}_2 \ldots\ \mathbf{O}_T$ uma sequência finita de observações onde cada observação é um símbolo discreto retirado de um alfabeto finito de símbolos: os índices de quantificação vectorial.

Dada uma sequência de observações **O** e o modelo $\mathbf{M} = (\pi, \mathbf{A}, \mathbf{B})$ com **N** estados, a probabilidade de **O** ter sido gerada por este modelo é[11]:

$$\mathbf{P}(\mathbf{O} \mid \mathbf{M}) = \sum_{i_1, i_2, \ldots, i_T} \pi_{i_1}\ \mathbf{b}_{i_1}(\mathbf{o}_1)\ \mathbf{a}_{i_1, i_2} \ldots\ \mathbf{a}_{i_{T-1}, i_T}\ \mathbf{b}_{i_T}(\mathbf{o}_T)$$

O significado da equação anterior é o seguinte: Inicialmente, $\mathbf{t} = \mathbf{o}$, estar-se-á no estado $\mathbf{i}_1$ com probabilidade π_{i_1}, e é "gerado" o símbolo $\mathbf{o}_1$ com probabilidade $\mathbf{b}_{i_1}(\mathbf{o}_1)$. Em seguida transitar-se-á para o estado seguinte, $\mathbf{i}_2$, com probabilidade $\mathbf{a}_{i_1 i_2}$, e gera-se o símbolo $\mathbf{o}_2$ com probabilidade $\mathbf{b}_{i_2}(\mathbf{o}_2)$. Este processo prosseguirá até se atingir o último estado.

O somatório anterior pode ser calculado de um modo mais eficiente se definirmos probabilidade progressiva, $\alpha_t(\mathbf{i})$:

$$\alpha_t(\mathbf{i}) \triangleq \mathbf{P}(\mathbf{o}_1\ \mathbf{o}_2 \ldots\ \mathbf{o}_t,\ \mathbf{q}_i \text{ em } \mathbf{t} \mid \mathbf{M})$$

isto é, a probabilidade de observar a sequência $\mathbf{o}_1\ \mathbf{o}_2 \ldots\ \mathbf{o}_{T-1}$, estando no estado $\mathbf{q}_i$ num dado instante **t**, sabendo ou dado um modelo **M**.

Pode-se, agora, obter a seguinte relação de recursividade:

$$\alpha_{t+1}(\mathbf{j}) = [\sum_{i=1}^{N} \alpha_t(\mathbf{i})\ \mathbf{a}_{ij}]\ \mathbf{b}_j(\mathbf{o}_{t+1}) \qquad \mathbf{t} = 1, 2, \ldots, \mathbf{T}-1$$

Definindo probabilidade regressiva:

$$\beta_t(\mathbf{j}) \triangleq \mathbf{P}(\mathbf{o}_{t+1}\ \mathbf{o}_{t+2} \ldots\ \mathbf{o}_T \mid \mathbf{q}_j \text{ em } \mathbf{t}, \mathbf{M})$$

como sendo a probabilidade de observar a sequência $\mathbf{o}_{t+1}\ \mathbf{o}_{t+2} \ldots\ \mathbf{o}_T$ sabendo que se ocupava o estado $\mathbf{q}_j$ em **t** e um dado modelo **M**, obteremos a seguinte relação de recursividade:

$$\beta_t(\mathbf{i}) = \sum_{j=1}^{N} \mathbf{a}_{ij}\ \mathbf{b}_j(\mathbf{o}_{t+1})\ \beta_{t+1}(\mathbf{j}) \qquad \mathbf{T}-1 \geqslant \mathbf{t} \geqslant 1$$

e poderemos então determinar a probabilidade de **P** na forma:

$$\mathbf{P} = \mathbf{P}(\mathbf{O} \mid \mathbf{M}) = \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_t(\mathbf{i})\ \mathbf{a}_{ij}\ \mathbf{b}_j(\mathbf{o}_{t+1})\ \beta_{t+1}(\mathbf{j}) \qquad 1 \leqslant t \leqslant T-1$$

ou ainda, utilizando apenas probabilidades progressivas:

$$\mathbf{P} = \sum_{i=1}^{N} \alpha_T(\mathbf{i})$$

Ao utilizar a equação anterior para o cálculo de **P** está-se a incluir implicitamente todas as sequências de estados dentro de um determinado modelo que poderiam ter gerado a observação.

Poder-se-ia definir **P** como o máximo, para todas as sequências de estados $\mathbf{i} = \mathbf{i}_0, \mathbf{i}_1, \ldots, \mathbf{i}_T$ da probabilidade conjunta $\mathbf{P}(\mathbf{o}, \mathbf{i})$. Esta sequência de estados e a correspondente probabilidade da sequência de observação podem ser determinadas recorrendo ao algoritmo de Viterbi[11]:

$$\text{início:} \quad \phi_1(\mathbf{i}) = \pi_i\ \mathbf{b}_i(\mathbf{o}_1) \qquad 1 \leqslant \mathbf{i} \leqslant \mathbf{N}$$
$$\psi_1(\mathbf{i}) = \mathbf{0}$$

recursão: $\phi_t\ (j) = \max\limits_{1 \leqslant i \leqslant N} [\phi_{t-1}\ (i)\ a_{ij}]\ b_j\ (o_t)$

$$\psi_t\ (j) = \arg\max_{1 \leqslant i \leqslant N} [\phi_{t-1}\ (i)\ a_{ij}]$$

$$2 \leqslant t \leqslant T \qquad 1 \leqslant j \leqslant N$$

Este algoritmo, além de permitir determinar qual o modelo que poderia ter dado **O** como observação, determina ainda qual a melhor sequência de estados que origina $\mathbf{P}_{max}$:

finalização: $\mathbf{P}^* = \max\limits_{1 \leqslant i \leqslant N} [\phi_T\ (i)]$

$$i_T^* = \arg\max_{1 \leqslant i \leqslant N} [\phi_T\ (i)]$$

Para se obter a sequência de estados mais provável: para $t = T-1,\ T-2, \ldots,\ 1$:

$$i_t^* = \psi_{t+1}\ (i_{t+1}^*)$$

Repare-se que se está a caminhar do último estado para o primeiro estado da cadeia ("backtracking").

4.2 Treino do modelo

Sendo impossível obter um conjunto infinito de observações, o modelo que se irá obter não será óptimo pelo que normalmente, o modelo obtido corresponderá a um máximo local de **P**. Consideremos as probabilidades progressiva e regressiva, e determine-se o número expectável de transições γ_{ij} de $\mathbf{q}_i$ para $\mathbf{q}_j$, condicionada à sequência observada:

$$\gamma_{ij} = \frac{1}{P} \sum_{t=1}^{T-1} \alpha_t\ (i)\ a_{ij}\ b_j\ (o_{t+1})\ \beta_{t+1}\ (j)$$

Deste modo, o número expectável de transições do estado $\mathbf{q}_i$ é dado por γ_i:

$$\gamma_i = \frac{1}{P} \sum_{t=1}^{T-1} \alpha_t\ (i)\ \beta_t\ (i)$$

Recorde-se que $\gamma_i = \sum\limits_{j=1}^{N} \gamma_{ij}$

O quociente γ_{ij}/γ_i representa uma estimativa da probabilidade de nos encontrarmos no estado $\mathbf{q}_j$ sabendo que o anterior era $\mathbf{q}_i$. Este quociente é uma estimativa da probabilidade de transição $\mathbf{a}_{ij}$:

$$\bar{a}_{ij} = \frac{\gamma_{ij}}{\gamma_i} = \frac{\sum\limits_{t=1}^{T-1} \alpha_t\ (i)\ a_{ij}\ b_j\ (o_{t+1})\beta_{t+1}(j)}{\sum\limits_{t=1}^{T-1} \alpha_t\ (i)\ \beta_t\ (i)}$$

De modo semelhante se pode determinar uma estimativa de $\mathbf{b}_{jk}$ como a frequência de ocorrência de $\mathbf{v}_k$ no estado $\mathbf{q}_j$; teremos, em termos de α e de β:

$$\bar{b}_{jk} = \frac{\sum\limits_{t \ni o_t = v_k} \alpha_t\ (j)\ \beta_t\ (j)}{\sum\limits_{t=1}^{T} \alpha_t\ (j)\ \beta_t\ (j)}$$

E, finalmente, para a probabilidade inicial Π_i:

$$\bar{\Pi}_i = \frac{1}{P}\ \alpha_1\ (i)\ \beta_1\ (i)$$

Demonstra-se[12] que estas estimativas incrementam **P** excepto num ponto crítico que corresponde a máximos locais. As equações anteriores derivam do algoritmo de Baum-Welch (B.W.)[12].

4.3 Estruturas de HMM e número de estados

Consideram-se, pela sua generalidade, três tipos de estruturas: não restringidas e restringidas em cascata e paralelo, fig. 6:

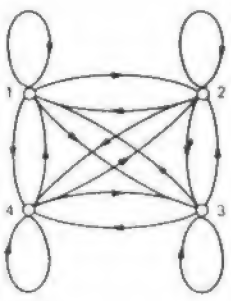


Fig. 6a) Estruturas de cadeias de Markov, não restringidas

Fig. 6b) Estruturas de cadeias de Markov, restringidas, em cascata

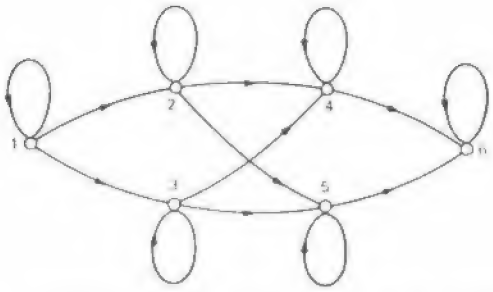


Fig. 6c) Estruturas de cadeias de Markov, restringidas, em paralelo

Os modelos com as restrições apresentadas têm matrizes de transição triangulares superiores; também denominados modelos esquerda-direita. Os modelos de estrutura em cascata podem progredir sequencialmente da esquerda para a direita, sendo possível o salto de um ou mais estados; os de estrutura em paralelo permitem caminhos múltiplos ao longo do modelo, continuando a evoluir da esquerda para a direita.

Qualquer das estruturas anteriores pode apresentar um número de estados arbitrário; notemos, contudo, que o número de parâmetros evolui com N^2 (matriz **A**), **N*M** (matriz **B**), e se é utilizado um número de estados grande ($\geqslant 10$) o cálculo de **P** vem fortemente penalizado. Repare-se que o modo de teste é o mais utilizado, pelo que os cálculos deverão ser rápidos; quanto ao treino do modelo, pode correr fora do período de utilização do sistema, e ocorrerá, geralmente, antes de o utilizar pela primeira vez, e eventualmente, para posteriores actualizações do vocabulário. Aparentemente não há um processo simples de determinar quantos estados deverá ter o modelo, uma vez que os estados de cadeia de Markov não têm que estar relacionados fisicamente com o fenómeno em observação.

Rabiner et al[11] para o reconhecimento dos dez dígitos determinaram a evolução do erro de detecção para um número de estados variável de 2 a 20 em estruturas em cascata, fig. 7:

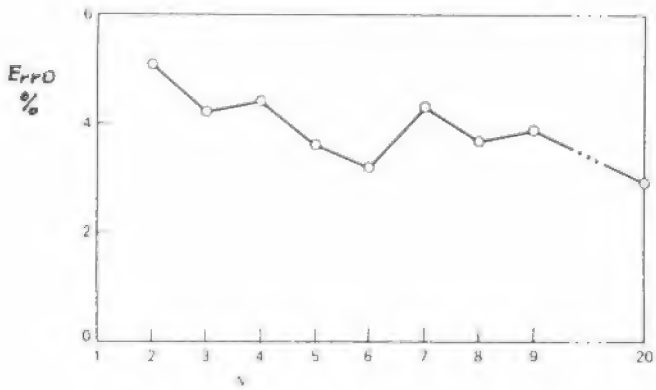


Fig. 7 Evolução do erro com N

Nota-se uma diminuição do erro com o aumento do número de estados, contudo, esse decremento é menor a partir de 8-9 estados, pelo que o acréscimos de complexidade do modelo não é justificado pelos resultados obtidos. As flutuações do erro em torno do valor médio são devidas à escolha inicial das matrizes **A** e **B**.

Na figura seguinte, são representadas as evoluções do erro de reconhecimento para cada um dos dígitos em função do número de estados, e mais uma vez se verifica que para este exemplo concreto, 5 ou 6 é o número de estados que optimiza o erro de reconhecimento bem como a velocidade de processamento[11]. Note-se ainda que para palavras monossilábicas ("one, two") o erro é maior do que para palavras bissilábicas ("zero, seven"), o que leva a crer que não existe relação simples com o número de estados e a duração de sons elementares[11].

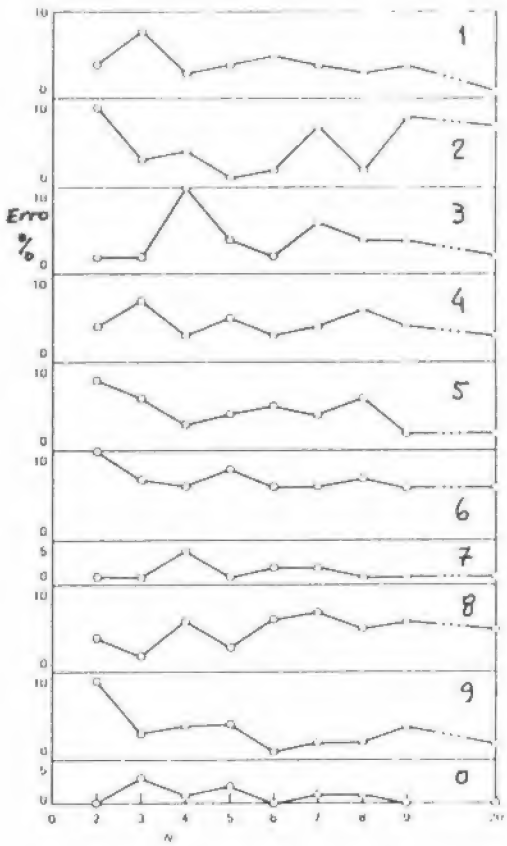


Fig. 8 Evolução do erro com N para cada um dos dígitos

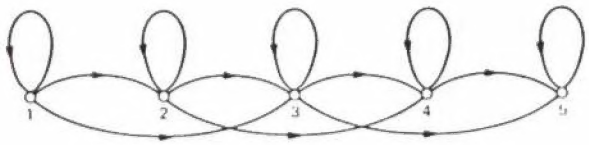
4.4 Sequências múltiplas de observação

Numa tentativa de aumentar o desempenho do sistema de reconhecimento quer para sistemas mono-utilizador que para multi-utilizador, é acessório ter várias observações de uma mesma palavra de modo a caracterizar as suas diferentes realizações. Tal facto irá penalizar bastante o processo de aprendizagem. Realizações práticas

utilizam de 2 a 10 modelos para a mesma palavra a ser reconhecida, o que pode implicar um aumento do treino de cerca de uma ordem de grandeza. É claro que o teste ou classificação também é afectado, contudo, o desempenho superior pode compensar o acréscimo de processamento, o que dependerá da aplicação em vista.

4.5 Restrições durante o treino

Para estruturas não restringidas, os elementos das matrizes **A** e **B** podem assumir quaisquer valores; nas restringidas, alguns desses elementos serão forçosamente nulos. Assim e para os modelos da fig. 9 teremos:



$a_{ij} = 0 \quad j<i \quad j \geqslant i+3$

$a_{ij} = 0 \quad j<i \quad j \geqslant i+2$

Fig. 9 Modelos de Markov restringidos

Quanto à matriz **B**, e pelo facto de as estimativas iniciais poderem originar algum $b_j(k) = 0$, pode suceder que $\alpha_{t-1}(i).a_{ij}$ seja diferente de zero para apenas um valor de **j**, e como tal a probabilidade dessa observação $(0_t = k)$ provir do modelo é nula — o que corresponde a um erro de falta de aprendizagem ou aprendizagem insuficiente. Um modo de resolver esta questão consiste em não anular valores de $b_j(k)$ cuja estimativa dê nula, ou seja impor $b_j(k) \geqslant \epsilon$ sendo ϵ um limiar escolhido. Teremos que os $b_j(k)$ inferiores ao limiar terão o seu valor modificado para ϵ, os $b_j(k)$ superiores a ϵ são escalados de modo a manter $\sum_k b_j(k) = 1$.

Na figura seguinte, fig. 10 o erro de reconhecimento é obtido para dois modelos com estruturas distintas[11]: um é não restringido, e o outro é-o, ambos têm 5 estados. A variável independente é o limiar ϵ que varia de 0 a 10^{-3}, verificando-se um erro de reconhecimento mais elevado, o que confirma o atrás exposto. Uma vez que o modelo não restringido tem um desempenho inferior serão utilizados apenas modelos restringidos, e destes os em cascata. Para $b_j(k)$ qualquer valor não perto de zero é aceitável, e como exemplo, poderemos ter $b_j(k) = 10^{-5}$.

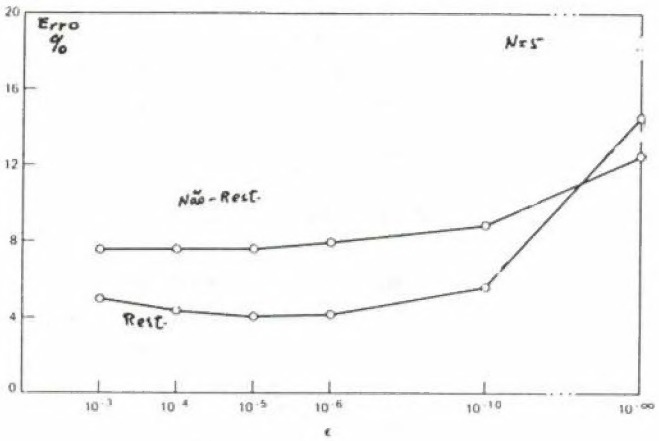


Fig. 10 Erro de reconhecimento, modelos restringido e não restringido, 5 estados

4.6 Escalamento

O método de aprendizagem apresentado necessita do cálculo de α e de β para $1 \leqslant t \leqslant T$, $1 \leqslant i \leqslant N$, contudo, das equações que nos permitem determinar estas probabilidades verifica-se que elas tendem exponencialmente para zero quando **T** ∞. Deste modo, para se treinar convenientemente um modelo, ocorrerá uma situação de "underflow" que pode ser resolvida mediante um escalamento conveniente de α e de β. Baseia-se este escalamento no produto de α por um coeficiente independente do estado de modo a que esta permaneça dentro da gama dinâmica do computador; β sofrerá um escalamento idêntico. Para uma descrição mais pormenorizada do escalamento das probabilidades progressiva e regressiva, veja-se[12].

4.7 Modelos de Markov esquerda-direita

O uso de modelos que evoluam sequencialmente da esquerda para a direita tem-se revelado muito útil nos sistemas de reconhecimento de fala. Apresentam as seguintes propriedades;

1) a primeira observação é produzida quando o modelo se encontra num determinado estado, denominado estado inicial, q_1.
2) a última observação é produzida com o modelo num estado final, de onde não transitará a não ser para ele próprio (absorvente), q_N.
3) uma vez que um estado seja abandonado, não será mais ocupado.

Nestas condições teremos:

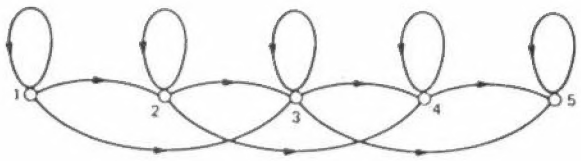


Fig. 11 Modelo de Markov esquerda-direita

No caso de se utilizar um modelo nestes pressupostos, para a sua aprendizagem não interessa uma sequência de observação muito longa, porque uma vez atingido o estado final o que resta da sequência de observações não traz nenhuma informação sobre os estados anteriores.

O conjunto de treino para este tipo de modelos consistirá num conjunto de observações obtidas por diferentes começos no estado inicial. Se se tratar do reconhecimento de palavras isoladas, várias fonações da mesma palavra constituirão o conjunto de treino. Deste modo, e para calcular a probabilidade de uma sequência de observações $\mathbf{o}_1, \mathbf{o}_2, \ldots, \mathbf{o}_T$ ter sido gerada pelo modelo, sabendo ainda que (e de acordo com as propriedades 1 a 3) que $\mathbf{o}_1$ e $\mathbf{o}_T$ teriam sido produzidas nos estados inicial e final, respectivamente:

$$\Pi = (1, 0, \ldots, 0)$$

$$B_T(j) = 1 \qquad 1 \leqslant j \leqslant N$$

Quanto à re-estimativa de $\mathbf{a}_{ij}$ para satisfazer a condição 3, basta impor $\mathbf{a}_{ij} = 0$ para $j<i$ e para outras situações que se entenda manter (por exemplo, o número máximo de estados a saltar). Esta condição é suficiente para manter as re-estimativas das probabilidades de transição nulas se se utilizar o algoritmo de Baum-Welch. Contudo, este algoritmo terá de ser adaptado às múltiplas observações da mesma palavra, pelo que teremos:

Seja $O = \{O^{(1)}, O^{(2)}, \ldots, O^{(k)}\}$ o conjunto de sequências de treino em que $O^{(k)}$ é a sequência índice **k**.

Todas as observações são supostas serem independentes entre si, e ajustar-se-ão os parâmetros do modelo **M** para maximizar:

$$P = \prod_{k=1}^{K} P(O^{(k)} \mid M) = \prod_{k=1}^{K} P_k$$

Uma vez que o algoritmo de B.W. determina a frequência de ocorrência de vários acontecimentos, é apenas necessário calcular cada uma dessas frequências relativas em cada sequência separadamente e adicioná-las para todas as sequências. Teremos assim:

$$\bar{a}_{ij} = \frac{\sum_{k=1}^{K} \sum_{t=1}^{T_{k-1}} \alpha_t^{(k)}(i)\, a_{ij}\, b_j(O_{t+1}^{(k)})\, \beta_{t+1}^{(k)}(j)}{\sum_{k=1}^{K} \sum_{t=1}^{T_{k-1}} \alpha_t^{(k)}(i)\, \beta_t^{(k)}(i)}$$

$$\bar{b}_{ij} = \frac{\sum_{k=1}^{K} \sum_{t \ni O_t^{(k)} = v_j} \alpha_t^{(k)}(i)\, \beta_t^{(k)}(i)}{\sum_{k=1}^{K} \sum_{t=1}^{T_k} \alpha_t^{(k)}(i)\, \beta_t^{(k)}(i)}$$

Note-se que o vector Π_i não é reestimado.

6. CONCLUSÕES

Como se referiu anteriormente, o algoritmo de B.W. não garante a obtenção de máximos globais de **P(O/M)** mas apenas de máximos locais. Contudo, e por inspecção de **P(O/M)** pode-se determinar, para um modelo esquerda-direita, a sequência de estados mais provável de comprimento **T** e que será:

$$S_{i(1)}, S_{i(2)}, \ldots, (T-N) \mathbin{@} S_{i(n^*)}, \ldots, S_{i(N)} \text{ com } T>N$$

sendo **R @ S_i** a sequência de **R** repetições do estado **i**.

Esta sequência pode ser fornecida como condição inicial para o treino do modelo e conduzirá com maior probabilidade a um modelo que estará mais perto do óptimo do que um obtido com uma condição inicial aleatória. A escolha de um modelo esquerdo-direita para modelo a utilizar tem a sua razão assente quer num menor número de cálculos necessários para o treino do modelo (restrições na matriz **A**), quer numa tentativa de fazer corresponder a uma evolução temporal da fala numa evolução do modelo da esquerda para a direita.

Os modelos de Markov aplicados ao reconhecimento de palavras isoladas têm dado bons resultados quer em valor absoluto: desempenhos superiores a 90%, quer quando comparados com outros métodos de reconhecimento, nomeadamente o DTW, especialmente se este também incluir quantificações vectorial dos parâmetros. Se os HMM utilizarem como saídas dos seus estados processos aleatórios multivariáveis com densidades de probabilidade conjunta contínuas, o seu desempenho é idêntico ao dos DTW. A vantagem da utilização dos HMM reside ainda numa maior rapidez do processo de classificação em que é francamente superior ao DTW; contudo, o treino é, do ponto de vista do tempo necessário mais demorado relativamente ao DTW. Para aplicações correntes este limitação não determinante.

Resta referir alguns dados sobre as prestações típicas de reconhecimento baseadas no DTW e no HMM; assim, e de acordo com resultados obtidos nos Laboratórios Bell, teremos que, para reconhecer um entre os dez dígitos se obtém uma taxa de sucessos de 98,2% para o DTW e de 97,1% para o HMM este último com quantificação vectorial (dimensão do codebook: 128 codewords). Estes valores dizem respeito a sistemas multi-utilizador.

BIBLIOGRAFIA

1. Wayne Lea, "Trends in Speech Recognition". Prentice-Hall, Inc., (1980).
2. Athanasios Papoulis, "Probability, Random Variables, and Stochastic Processes". ISE, McGraw-Hill (1984).
3. Lawrence Rabiner, Stephen Levinson, "Isolated and Connected Word Recognition — Theory and Selected Applications". IEEE Trans. Comm. (Maio 1981).
4. Andrés Buzo, Augustine Gray, Robert Gray, John Markel, "Speech Coding Based Upon Vector Quantization". IEEE, Trans. Acous Spe. Sig. Proc. (Outubro 1980).
5. Juang, "On the Hidden Markov Model and Dynamic Time Warping for Speech Recognition — A Unified View". BSTJ (Setembro 1984).
6. Alan Poritz, "Linear Predictive Hidden Markov Models and the Speech Signal". ICASSP (1982).
7. Louis Liporace, "Maximum Likelihood Estimation for Multivariate Observations of Markov Sources". IEEE, Trans. Infor. Theory (Setembro 1982).
8. L. Rabiner, S. Levinson, M. Sondhi, "On the Use of Hidden Markov Models for Speaker-Independent Recognition of Isolated Words from a Medium — Size Vocabulary". BSTJ (Abril 1984).
9. S. Levinson, L. Rabiner, M. Sondhi, "Speaker Independent Isolated Digit Recognition Using Hidden Markov Models". ICASSP (1983).
10. Robert Billi, "Vector Quantization and Markov Source Models Applied to Speach Recognition". ICASSP (1982).
11. L. Rabiner, S. Levinson, M. Sondhi, "On the Application of Vector Quantization and Hidden Markov Models to Speaker — Independent, Isolated Word Recognition". BSTJ (Abril 1983).
12. Levinson, L. Rabiner, M. Sondhi, "An Introduction to the Application of the Theory of Probabilistic Functions of a Markov Process to Automatic Speech Recognition". BSTJ (Abril 1983).
13. M. Russel, R. Moore, "Explicit Modelling os State Occupancy in Hidden Markov Models for Automatic Speech Redcognition". ICASSP (1985).
14. B. Huang, L. Rabiner, S. Levinson, M. Sondhi, "Recent Developments in the Application of Hidden Markov Models to Speaker — Independent Isolated Word Recognition". ICASSP (1985).
15. A. Newel, J. Barnett, J.W. Forgie, C. Green, D. Klatt, J. Licklider, J. Munson, D. Reddy, W. Woods, "Speech Understanding Systems". North-Holland/American Elsevier (1973).
16. C. Myers, L. Rabiner, A. Rosenberg, "Performance Tradeoffs in Dynamic Time Warping Algorithms for Isolated Word Recognition". IEEE, ASSP (Dezembro 1980).
17. L. Rabiner, A. Rosenberg, S. Levinson, "Considerations in Dynamic Time Warping Algorithms for Discrete Word Recognition". IEEE, ASSP (Dezembro 1978).
18. L. Rabiner, "Pattern Recognition Algorithms for Isolated and Connected Word Recognition". ICASSP (1985).
19. J. Tribolet, L. Rabiner, J. Wilpon. "An Improved Model for Isolated Word Recognition". BSTJ (Novembro de 1982).
20. "Voice Recognition Module, Model VRQ499", Interstate Electronics Corporation, Voice Products Operations Data Sheet (1981).

A ESTABILIDADE ESTRUTURAL E O COMPORTAMENTO POS-ENCURVADURA

A.J. REIS

Por lapso verificado na montagem do último número da Revista Técnica, não foi inserido no artigo A ESTABILIDADE ESTRUTURAL E O COMPORTAMENTO PÓS-ENCURVADURA *da autoria do Prof. A. J. Reis, o quadro que agora reproduzimos. Pelo facto apresentamos as nossas desculpas.*

QUADRO I

Ponte sobre o Danúbio em Viena colapso em 6-NOV.-1969

Banzo comprimido do caixão — Placa de 7600 × 10 mm com 12 reforços de 160 × 12 mm;
$\sigma_c = 244\ N/mm^2$

MÉTODO DE CÁLCULO	IMPERFEIÇÕES δ_o (mm) + σ_{oc} (N mm^{-2})	TENSÃO MÉDIA DE COLAPSO LOCAL σ_m (N mm^{-2})
VON KARMAN	0	254
KOITER	0	255
WALKER & MURRAY	$\alpha = 0.3$ ($\delta_o = 3.4$)	186
	$\alpha = 0.4$ ($\delta_o = 4.5$)	168
HORNE & NARAYANAN	$\delta_o = 3$ ($\delta'_o = 11$)	182
	$\delta_o = 5.8$ ($\delta'_o = 17$)	165
ECCS	$c^* = 0.83$	184
FRIEZE	$\delta_o = 3$; $\sigma_{oc} = 22$	180

COLECÇÃO DE PROGRAMAS E FILTROS DESTINADOS A OPERAR SOBRE "CONJUNTOS VAGOS"

A. J. N. Martins Aguiar*

RESUMO

Colecção compreensiva de programas aplicáveis a conjuntos vagos (Fuzzy sets) e respectivas demonstrações.

ABSTRACT

Comprehensive collection of programs applicable to Fuzzy sets and respective proofs.

Apresentação

O objectivo da apresentação deste programa e respectiva justificação, é pôr à disposição do utente um conjunto de «filtros» bem como das conectivas usadas na Teoria dos Conjuntos Vagos (Fuzzy Sets, Ensembles Flous, Conjuntos Borrosos).

Na verdade, sem estas subrotinas torna-se pesado operar no domínio dos conjuntos vagos.

Este trabalho além de resumir alguns trabalhos anteriores realizados nas cadeiras de Automação I e II, cujos nomes são referidos na Bibliografia, introduz todo um conjunto de demonstrações formais e de novas subrotinas.

Toda esta matéria foi programada em Fortran IV porque é uma linguagem relativamente geral e daí que susceptível de correr em computadores de diversos tipos, e é assim possível usar este conjunto de programas na resolução de problemas propostos em «Fuzzy Mathematics».

Outros programas destinados a completar as aplicações dos Conjuntos Vagos serão apresentados em futuros artigos.

Os fundamentos teóricos estão referidos na Bibliografia.

A. G. PORTELA **

Introdução

As conectivas definidas nos conjuntos vagos são apoiadas nas conectivas definidas nos reticulados onde se descrevem os respectivos «caracterizantes».

Assim é essencial por um lado identificar os reticulados referidos para escolher as conectivas aplicáveis, a fim de operar sobre os conjuntos vagos subjacentes.

Nesta ordem de ideias será apresentado:

1.º Capítulo

CARACTERIZAÇÃO DO CONJUNTO ONDE UMA RELAÇÃO DE ORDEM FOI IMPOSTA

1.1 — Introdução.
1.2 — Verificação das propriedades do grafo

a) Existência e unicidade de ínfimo e supremo do grafo;
b) Existência de circulações;
c) Existência de ramificações;
d) Resumo e conclusão de 1.2.

* Assistente Estagiário do Instituto Superior Técnico.
** Professor Catedrático do Instituto Superior Técnico.
Original recebido para publicação em 8/5/85.